

Labeled Peptides, Proteins and Antibodies and Processes and Intermediates
Useful for Their Preparation

5 **Government Funding**

The invention described herein was made with United States Government support under Grant Number MCB-9812248 awarded by the National Science Foundation, and under Grant Numbers CA58689, AG15430 and GM 57464 awarded by the National Institutes of Health. The United States
10 Government has certain rights in this invention.

Background of the Invention

Modified peptides and proteins are valuable biophysical tools for studying biological processes, both *in vitro* and *in vivo*. They are also useful in
15 assays to identify new drugs and therapeutic agents. In particular, quantitative live cell imaging using fluorescent proteins and peptides is revolutionizing the study of cell biology. An exciting recent development within this field has been the construction of peptide and protein biosensors exhibiting altered fluorescence properties in response to changes in their environment, oligomeric state,
20 conformation upon ligand binding, structure, or direct ligand binding. Appropriately labeled fluorescent biomolecules allow spatial and temporal detection of biochemical reactions inside living cells. See for example Giuliano, K.A., et al., *Annu. Rev. Biophys. Biomol. Struct.* 1995, 24:405-434; Day, R.N. *Mol. Endocrinol.* 1998, 12:1410-9; Adams, S.R., et al., *Nature* 1991, 349:694;
25 Miyawaski, A., et al., *Nature* 1997, 388:882-7; Hahn, K., et al., *Nature* 1992, 359:736; Hahn, K.M., et al., *J. Biol. Chem.* 1990, 265:20335; and Richieri, G.V., et al., *Mol. Cell. Biochem.* 1999, 192:87-94.

Procedures for site-specific modification of polypeptides have been described, including: chemically selective labeling in solution (Brinkley, M.
30 *Bioconjugate Chemistry* 1992, 3:2-13) and on resin bound peptides (Hackeng, T., et al., *J. Biol. Chem.* Submitted); introduction of ketone amino acids through synthetic procedures (Rose, K. *J. Am. Chem. Soc.* 1994, 116:30-33; King, T.P., et al., *Biochemistry*, 1986, 25:5774-5779; Rose, K., et al., *Bioconjugate Chem.* 1996, 7:552-556; Marcaurelle, L.A., Bertozzi, C.R. *Tett. Lett.* 1998, 39:7279-

7282; and Wahl, F., Mutter, M. *Tett. Lett.* 1996, 37:6861-6864); and molecular biology techniques (Cornish, V.W., et al., *J. Am. Chem. Soc.* 1996, 118:8150).

While each of these methods has utility for producing a particular class of biosensor or labeled polypeptide, all have limitations that restrict their general use. Labeling of natural amino acid side-chains in solution is often impractical because of the existence of many other competing nucleophiles. Additionally, the use of unnatural amino acids, such as those bearing ketones for selective labeling, requires the synthesis of dye constructs or amino acids that are difficult to make and not available commercially.

Currently, the major obstacles to the development of fluorescent biosensors remain: (1) The difficulty in site-specific placement of the dye in the polypeptide and (2) determining exactly which site is optimal for dye placement (Giuliano, K.A., et al., *Annu. Rev. Biophys. Biomol. Struct.* 1995, 24:405-434). Solvent-sensitive dyes and other biophysical probes must be placed precisely for optimal response to changes in protein structure without interference with biological activity. Also, the need for site-specific incorporation of two dyes without impairment of biological activity has proven a serious limitation for utilization of fluorescence resonance energy transfer (FRET) within a single protein. Total chemical synthesis of proteins provides a potential solution to these problems (Wilken, J., Kent, S.B.H. *Curr. Op. Biotechnology.* 1998, 9:412; Kent, S.B.H. *Ann. Rev. Biochem.* 1988, 57, 957-989; Dawson, P.E., et al., *Science* 1994, 266:776-779; Muir, T.W., et al., *Proc. Natl. Acad. Sci.* 1998, 95:6705-6710; and Cotton, G.J., et al., *J. Am. Chem. Soc.* 1999, 121:1100-1101). However, many biophysical probes suitable for fluorescent biosensors or other purposes are not stable to the various conditions used for peptide synthesis, and site-specific incorporation after synthesis has been difficult to achieve.

Moreover, labeling with hydrophobic dyes such as thionine or methylene blue can be problematic because these dyes autoaggregate in aqueous solution at high concentration. See, *J. Am. Chem. Soc.* 63, 69 (1941). These aggregates cause a change in the absorption spectrum and a reduction in the fluorescence of the dyes. Cyanines and merocyanines are also thought to aggregate, causing a quenching of fluorescence (*J. Phys. Chem.* 69, 1894 (1965)). Such aggregation interferes with conjugation of these fluorescent dyes to other molecules such as proteins. Moreover aggregation by cyanines and merocyanines can be

exacerbated after the dyes are conjugated. For example, Waggoner et al. have observed an aggregation phenomenon following the conjugation of cyanin isothiocyanate with an antibody (Cytometry 10, 11-19 (1989)). Fluorescence of a conjugate between a cyanin fluorescent dye and an anti-HCG antibody (molar ratio=1.7) named CY5.18 is quenched in comparison with that of the free cyanin (see U.S. Pat. No. 5,268,486 and Anal. Biochem. 217, 197-204 (1994)). Also, while cyanines are generally stable, inexpensive, simple to conjugate to other molecules and of a suitable size for the recognition of small molecules, they do not change their fluorescence in response to environmental factors, such as solvent polarity. New dyes that eliminate these problems are needed.

Thus, there is currently a need for new fluorescent dyes and peptide synthons having protected functional groups that can be selectively modified to incorporate one or more functional molecules (e.g. a fluorescent label) following peptide synthesis. There is also a need for proteins and antibodies with biophysical probes attached to precise locations, and for simple, non-destructive methods of making such labeled proteins and antibodies. Simpler methods for using these labeled peptides, proteins and antibodies in vivo as biosensors are also needed.

Summary of the Invention

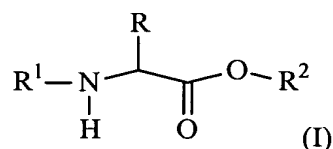
The present invention provides a highly efficient method for the site-specific attachment of biophysical probes or other molecules to unprotected peptides following chemical synthesis. The methodology utilizes amino acids having one or more protected aminooxy groups, which can be incorporated during solid-phase peptide synthesis or which can be combined with recombinant peptides through post expression steps. It has been discovered that the protected aminooxy group can be unmasked following peptide synthesis, and reacted with an electrophilic reagent to provide a modified (e.g. a labeled) peptide. The aminooxy group reacts selectively with electrophiles (e.g. an activated carboxylic ester such as an N-hydroxy-succinimide ester) in the presence of other nucleophilic groups including cysteine, lysine and amino groups.

Thus, selective peptide modification (e.g. labeling) can be accomplished after synthesis using commercially available and/or chemically sensitive

molecules (e.g. probes). The methodology is compatible with the synthesis of C α -thioester containing peptides and amide-forming ligations, required steps for the synthesis of proteins by either total chemical synthesis or expressed protein ligation. An aminooxy containing amino acid can be introduced into different sites by parallel peptide synthesis to generate a polypeptide analogue family with each member possessing a single specifically-labeled site. The parallel synthesis enables the development of optimized biosensors or other modified polypeptides through combinatorial screening of different attachment sites for maximal response and minimal perturbation of desired biological activity.

Thus, a simple and efficient methodology for site-specific modification (e.g. labeling) of peptides after synthesis has been developed that provides high yield, selectivity, and compatibility with both solid-phase peptide synthesis and C $^{\alpha}$ -thioester peptide recombinant synthesis.

Accordingly, the invention provides a synthetic intermediate (i.e. a synthon) useful for preparing modified peptides, which is a compound of formula (I):



wherein:

R¹ is hydrogen or an amino protecting group;

R² is hydrogen or a carboxy protecting group;

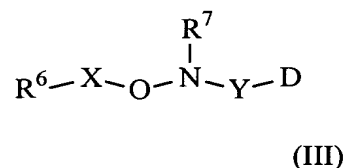
R is an organic radical comprising one or more aminooxy

Peptides including one or more aminooxy groups are also useful synthetic intermediates that can be modified to provide related peptides having altered biological, chemical, or physical properties, such as, for example, a peptide linked to a fluorescent label. Accordingly, the invention also provides a peptide having one or more (e.g. 1, 2, 3, or 4) aminooxy groups; provided the peptide is not glutathione. The invention also provides a peptide having one or more (e.g. 1, 2, 3, or 4) secondary aminooxy groups.

The invention generally provides intermediates and methods that allow for site-specific modification of peptides after synthesis. Accordingly,

functional molecules can be selectively linked to a peptide to provide a peptide conjugate having altered biological, chemical, or physical properties. For example, functional molecules (e.g. biophysical probes, peptides, polynucleotides, and therapeutic agents) can be linked to a peptide to provide a peptide conjugate having differing and useful properties.

Thus, the invention also provides a compound of formula (III):



wherein:

- R⁶ is a peptide, polypeptide or antibody;
 - X is a direct bond or a linking group;
 - R⁷ is hydrogen, (C₁-C₆)alkyl, an amino protecting group, or a radical comprising one or more aminooxy groups;
 - Y is a direct bond or a linking group; and
 - D is a functional molecule.
- Preferably the functional molecule is a biophysical probe, such as a fluorescent group that can be used for FRET studies or other studies involving fluorescent signals, such as excimer pair formation.

Processes for preparing synthons of the invention as well as the polypeptide, antibody and protein conjugates of the invention are provided as further embodiments of the invention and are illustrated by the procedures in the Examples below.

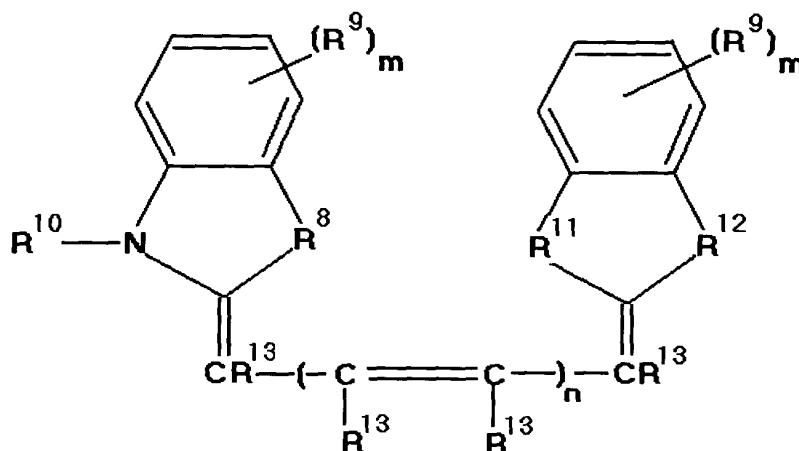
Thus, the invention also provides a method for preparing a peptide conjugate comprising a peptide and a functional molecule, comprising reacting a peptide having one or more aminooxy groups with a corresponding functional molecule having an electrophilic moiety, to provide the peptide conjugate.

The present invention further provides environment-sensing dyes that can be readily conjugated to proteins and other molecules without the problems of aggregation, fluorescence quenching and the like. The present dyes possess bright fluorescence and environmentally-sensitive fluorescence changes suitable for use in living cells. Unlike cyanine dyes, the environmental sensitivity of

these dyes can be used to form biosensors that report many aspects of protein behavior, or the behavior of other molecules. Protein behaviors including conformational change, phosphorylation state, ligand interaction, protein-protein binding and various post-translational modifications affect the distribution of charged and hydrophobic residues. The dyes can report these changes with changes in their fluorescence.

The present invention overcomes the disadvantages of the available environmentally-sensitive fluorescent dyes. The present fluorescent probes exhibit high fluorescence levels before and after conjugation to other molecules, including proteins and antibodies, and changes in fluorescence suitable for many purposes, including in vivo and in vitro assays of protein behavior.

The present invention provides new fluorescent dyes that can be used in any manner chosen by one of skill in the art. The dyes can be linked to any useful molecule known to one of skill in the art using any available procedure. In one embodiment the fluorescent dyes are linked to peptides, polypeptides or antibodies using the methods provided herein. These dyes have the following structure (IV).



wherein:

each m is separately an integer ranging from 1-3;

n is an integer ranging from 0 to 5;

R^8 , R^{11} and R^{12} are separately CO , SO_2 , $\text{C}=\text{C}(\text{CN})_2$, S , O or

$\text{C}(\text{CH}_3)_2$;

each R^{13} is alkyl, branched alkyl or heterocyclic ring derivatized with charged groups to enhance water solubility and enhance photostability;

- R^9 and R^{10} are chains carrying charged groups to enhance water solubility (i.e. sulfonate, amide, ether) and/or chains bearing reactive groups for conjugation to other molecules. The reactive group is a functional group that is chemically reactive (or that can be made chemically reactive) with functional groups typically found in biological materials, or functional groups that can be readily converted to chemically reactive derivatives using methods well known in the art. In one embodiment of the invention, the charged and reactive groups are separately haloacetamide ($-\text{NH}-(\text{C}=\text{O})-\text{CH}_2-\text{X}$), where X is Cl, Br or I. Alternatively, the charged and reactive groups are separately amine, maleimide, isocyanato ($-\text{N}=\text{C}=\text{O}$), isothiocyanato ($-\text{N}=\text{C}=\text{S}$), acyl halide, succinimidyl ester, or sulfosuccinimidyl ester. In another embodiment, the charged and reactive groups are carboxylic acid (COOH), or derivatives of a carboxylic acid.
- An appropriate derivative of a carboxylic acid includes an alkali or alkaline earth metal salt of carboxylic acid. Alternatively, the charged and reactive groups are reactive derivatives of a carboxylic acid ($-\text{COORx}$), where the reactive group Rx is one that activates the carbonyl group of $-\text{COORx}$ toward nucleophilic displacement. In particular, Rx is any group that activates the carbonyl towards nucleophilic displacement without being incorporated into the final displacement product. Examples of COORx : ester of phenol or naphthol that is further substituted by at least one strong electron withdrawing group, or carboxylic acid activated by carbodiimide, or acyl chloride, or succinimidyl or sulfosuccinimidyl ester. Additional charged and reactive groups include, among others, sulfonyl halides, sulfonyl azides, alcohols, thiols, semicarbazides, hydrazines or hydroxylamines.

- The invention still further provides a method of identifying an optimal position for placement of a functional molecule on a peptide having a peptide backbone and a known activity, which includes making a series of peptide conjugates, each peptide conjugate having the same amino acid sequence and the same functional molecule, wherein the functional molecule is linked at a different location along the backbone of every peptide conjugate in the series, and observing which functional group location does not substantially interfere with the known activity of the peptide.

The invention also provides a method of identifying an optimal position for placement of a functional molecule in a protein having a known activity and an identified peptide segment for attachment of the functional molecule, which includes making a series of peptide conjugates, each peptide conjugate having the amino acid sequence of the identified peptide segment and the same functional molecule, wherein the functional molecule is linked at a different location along the backbone of every peptide conjugate in the series; replacing the identified peptide segment in each protein of a series of said proteins with a peptide conjugate selected from the series of peptide conjugates to create a series of protein conjugates each having the functional group at a different location; and observing which functional group location does not substantially interfere with the known activity of the protein.

The invention further provides a method of identifying an optimal position for placement of an environmentally-sensitive functional molecule on a peptide biosensor having a backbone, which includes making a series of peptide conjugates, each peptide conjugate having the same amino acid sequence and the same functional molecule, wherein the functional molecule is at a different location along the backbone of every peptide conjugate in the series, and observing which functional group location provides the strongest signal change in response to an environmental change in the peptide conjugate. The signal change can be any observable change in a signal, for example, the change can be a change in fluorescence emission intensity, fluorescence duration or fluorescence emission wavelength. The environmental change in the peptide biosensor can be, for example, an interaction with a target molecule.

The invention still further provides a method of identifying an optimal position for placement of an environmentally-sensitive functional molecule in a protein having a known activity and an identified peptide segment for attachment of the functional molecule, which includes making a series of peptide conjugates, each peptide conjugate having the amino acid sequence of the identified peptide segment and the same environmentally-sensitive functional molecule, wherein the environmentally-sensitive functional molecule is linked at a different location along the backbone of every peptide conjugate in the series; replacing the identified peptide segment in each protein of a series of said proteins with a peptide conjugate selected from the series of peptide conjugates

to create a series of protein conjugates, each having the environmentally-sensitive functional group at a different location; and observing which functional group location provides the strongest signal change in response to an environmental change in the protein conjugate.

5 The invention also provides a method for detecting GTP activation of a Rho GTPase protein, which includes contacting a polypeptide biosensor with a test substance, wherein said polypeptide biosensor comprises a polypeptide capable of binding a GTP-activated Rho GTPase protein, and wherein said polypeptide is operatively linked to an environmentally sensitive fluorescent
10 dye; and observing fluorescence emissions from said polypeptide biosensor at a wavelength emitted by said fluorescent acceptor dye; wherein said environmentally sensitive fluorescent dye will emit light of a different intensity or a different wavelength when the polypeptide biosensor is bound to the GTP-activated Rho GTPase protein than when the polypeptide biosensor is not bound.

15

15
27
21

Brief Description of the Figures

FIG. 1 illustrates a general strategy for site-specific labeling of polypeptides. The protected aminooxy group is incorporated during solid-phase peptide synthesis (synthesis on a thioester-linker resin is shown); cleavage from
20 the resin generates a peptide possessing unprotected sidechains, an aminooxy group and a C-terminal thioester; and ligation and subsequent site-specific labeling produces the full-length peptide with a functional molecule attached at the aminooxy nitrogen.

FIG. 2 illustrates the synthesis of PA-test and SA-test peptides.

25 FIG. 3 shows HPLC analysis of purified SA-test peptide (top) and crude reaction products from optimized labeling conditions (bottom).

FIG. 4 illustrates the synthesis of a protected intermediate (4) of the invention.

FIG. 5 illustrates how biosensors can be employed according to the
30 present invention. In this example, the activation of Rac1 by GTP was observed. p21-Activated kinase (PAK) was used as a biosensor because PAK will only bind to Rac1 when Rac1 is activated by GTP. A Rac1-Green Fluorescent Protein (GFP-Rac1, shown as a green circle attached to a blue square) fusion protein was made and a cell line expressing this fusion protein

was generated. Cells expressing GFP-Rac1 were injected with a fragment of p21-activated kinase ("PBD") labeled with Alexa-546 dye (red circle). This fragment of PAK binds selectively to GFP-Rac-GTP, but not to GFP-Rac-GDP. Upon binding to GFP-Rac, the Alexa on the labeled fragment undergoes fluorescence resonance energy transfer (FRET) as the Alexa and GFP fluorophores are brought close together. This FRET can be measured within a living cell or in vitro to map the distribution, localization and level of Rac-GTP binding. FRET produces a unique fluorescence signal because excitation of GFP leads to emission from Alexa as energy is transferred from the excited GFP fluorophore to the nearby Alexa dye (J.R. Lakowicz, *Principles of Fluorescence Spectroscopy* (Plenum Press, New York, 1983), pp.305-341)). By imaging the cell with different wavelengths, both the distribution of Rac and Rac activation can be studied in the same cell. GFP excitation and emission are used for overall Rac distribution, while GFP excitation and Alexa emission are used for FRET.

FIG. 6 illustrates one method of using an environmentally sensitive fluorescent dye in the present methods so that changes in naturally existing proteins can be detected and observed *in vivo* or *in vitro*. Figure 6a depicts FRET between the Rac1-Green Fluorescent Protein (GFP-Rac1) fusion and the p21-activated kinase biosensor (PBD) labeled with Alexa-546 dye as shown in Figure 5. Inactive Rac1 is depicted as a gray circle with Green Fluorescent Protein (green circle) attached. Upon activation by GTP, Rac1 undergoes a structural change depicted as a gray circle changing to a half-rounded gray rectangle. Unbound PBD is depicted as a black L-shape with an attached Alexa-546 dye (orange open circle). Before Rac1 is activated, PBD cannot bind and the Alexa-546 cannot undergo FRET. However, after Rac1 activation, the Rac1 assumes a conformation that permits PAK binding. Such binding juxtaposes the Green Fluorescence Protein and the Alexa-546, which produces FRET. Figure 6b illustrates how an environmentally sensitive fluorescent dye eliminates the need to create a fusion protein like the GFP-Rac1 protein depicted in Figure 6a. In Figure 6b, a natural, unmodified protein is depicted as a gray oval. The protein changes conformation upon activation by GTP, which is depicted as the transition to a half-rounded gray rectangle. When the protein is in the activated state, a polypeptide that binds only to the activated state (black L-shape), with an attached environmentally sensitive dye (open orange circle), can bind. Upon

binding, the environmentally sensitive dye will emit light of a different wavelength, duration or intensity (filled orange circle) than before binding. Use of this type of environmentally sensitive dye is further illustrated in Figures 15-18.

5 FIG. 7 illustrates what conditions will optimally provide FRET between GFP-Rac and Alexa-PBD *in vitro*. Figure 7A shows fluorescence emission from solutions containing a fixed level of GFP-Rac bound to GTP γ S at different concentrations of Alexa-PBD (PBD labeled with Alexa-546). Light at 480 nm was selectively used for GFP excitation, and direct (non-FRET) excitation of
10 Alexa was subtracted from these spectra. In the absence of Alexa-PBD, the emission from GFP (peak at 508 nm) is maximal and no Alexa emission (peak at 568 nm) is seen. As the concentration of Alexa-PBD is increased, binding of Alexa-PBD to Rac-GFP leads to FRET, producing increasing emission at 568 nm and a decrease at 508 nm. Figure 8 shows variation of the 568/508 nm
15 emission ratios as a function of Alexa-PBD concentration for GFP-Rac bound to GTP γ S (circles) or to GDP (open squares). Figure 7B shows the variation of this same emission ratio with changes in the level of GTP-bound Rac. All data points were the average of three independent experiments.

 FIG. 8 illustrates how GFP-Rac expression levels and levels of
20 intracellular Alexa-PBD (as observed by fluorescence) correlate with changes in normal cell behavior produced by these proteins. Figure 8A shows what levels of GFP-Rac expression were correlated with ruffling. Cells with different expression levels of either wild type or constitutively active Q6 IL GFP-Rac (determined on the basis of GFP intensity /cell area) were scored for ruffling.
25 Each point represents an individual cell, placed in the higher (Ruffling) or lower (Nonruffling) row depending on whether ruffling was induced. As illustrated, there is a level of GFP intensity below which ruffling was consistently not induced by expression of wild type GFP-Rac. Only cells with Rac expression levels below 250 on this scale were used in biological experiments. The validity
30 of this approach was supported by scoring of constitutively active Rac, which showed ruffle induction at much lower levels of expression. Figure 8B shows which levels of intracellular Alexa-PBD would perturb normal serum-induced ruffling. Cells were scored as in Figure 8A, at different levels of Alexa-PBD fluorescence. Based on this experiment, only cells with Alexa-PBD

concentrations below 400 intensity units on this scale were used in biological experiments. These studies demonstrated that FRET could still be readily detected at appropriately low levels of introduced protein.

FIG. 9 illustrates the dynamics of Rac activation during growth factor stimulation of quiescent cells. Figure 9A provides photomicrographs showing Rac localization (GFP-Rac) and Rac activation (FRET) before stimulation of quiescent Swiss 3T3 fibroblasts. Figure 9B provides photomicrographs of the same Swiss 3T3 fibroblasts three minutes after addition of serum. Warmer colors correspond to higher intensity values. The cells showed accumulation of Rac at and around the nucleus before stimulation (GFP-Rac image). Most of the nuclear GFP-Rac was associated with the nuclear envelope. Serum or PDGF addition generated multiple moving ruffles that showed FRET, while no FRET was seen at the nucleus before or after stimulation. Of thirty-five cells stimulated with either serum or PDGF, thirty-one began ruffling within 15 minutes. FRET was seen in the ruffles of all but one of the ruffling cells. Figures 9C and 9D demonstrate that simple localization of Alexa-PBD is inferior to FRET in quantifying and localizing Rac-GTP binding (Bar = 8 μ m). The ruffle in Figure 9B is shown in close-up in Figure 9C, visualized using FRET. PBD localization in the same region is visualized in Figure 9D with scaling optimized for detection of the ruffle. Without prior knowledge of the ruffle's location, this localization would have been difficult to discern. The high background due to unbound PBD cannot be eliminated in Figure 9D and binding to other target proteins is not eliminated as it is in the highly specific FRET signal.

In each of the GFP-Rac images above, intensities range between 300-1100. The image of FRET before serum addition was scaled to demonstrate the low levels of FRET, with values ranging between 0 and 15. In the image of FRET after stimulation, the ruffle contains the highest values of 40 to 65. Nuclear FRET was not seen in any of the cells examined.

FIG. 10 illustrates Rac activation in motile cells. This figure shows two examples of the Rac 1 activation gradient, with high Rac1 activation at the leading edge of polarized, motile Swiss 3T3 fibroblasts (Bar = 24 μ m). In these experiments high levels of Rac-GTP were also frequently seen in the juxtannuclear region of the cell. The strong correlation of this gradient with the

direction of movement indicates that activated Rac is spatially organized in polarized cells to help guide or propagate movement. Comparison of the GFP and FRET images shows that the distribution of activated Rac does not parallel that of Rac itself. FRET intensities are 0-18 (top image) and 0-32 (bottom image). In the GFP images, intensities range between 98-700 (top image) and 100-1100 (bottom image).

FIG. 11 provides the structure of one of the present fluorescent dyes and the spectrum of fluorescence emission for that dye in water, methanol and butanol. As illustrated, the fluorescent emission of this dye increases with increasing solvent hydrophobicity.

FIG. 12 provides the structure of another fluorescent dye of the present invention and shows its spectrum in aqueous solution, compared to similar dyes lacking the groups designed to prevent aggregation. In the curve from each dye, the peak to the right is the unconjugated, highly fluorescent form of the dye, while that to the left is the weakly fluorescent form. The curve furthest to the right is the dye containing groups to prevent aggregation.

FIG. 13 provides one method for synthesizing a dye of the present invention. Conversion of compound 1 to an amine 2 followed by protection of the amine provides compound 3, which can be alkylated to give compound 4. Reaction of compound 4 with compound 9 provides compound 5, which can be deprotected to provide amine 6. Alkylation or acylation of amine 6 with a chain carrying a charged group, or with a chain bearing a reactive group for conjugation to another molecule, or with another molecule directly provides a dye of the invention 7. Intermediate compound 9 can be prepared by condensation of compound 8 with the requisite aldehyde, under conditions that are known in the art

FIG. 14 provides a three-dimensional image of the CRIB domain ("CBD") of the Wiscott-Aldrich syndrome protein (WASP) bound to Cdc42. The essential residues are depicted in yellow, hydrophobic residues depicted in purple and the red sites show positions where the new dyes were attached to generate changes in fluorescence when the CBD bound Cdc42.

FIG. 15 provides the intensity of fluorescence at various wavelengths observed when fluorescently labeled CBD binds to cdc42. When cdc42 is activated with GTP γ S, highly intense fluorescence is observed (brown line

labeled "GTP γ S"), compared to when no cdc42 is present (red line labeled "no cdc42"), or when cdc42 is not activated (orange line labeled "GDP").

FIG. 16 shows how the present methods can be used for in vitro assays on crude cellular lysates. In this case, the fluorescently labeled CBD was added to a neutrophil lysate. Upon binding to activated cdc42, CBD will emit
5 fluorescence of greater intensity. At time zero fMLP, which stimulates neutrophils to activate cdc42, was added to the cellular lysate and the amount of fluorescence generated by the lysate (●) was measured as a function of time. As
10 a control, the maximal amount of fluorescence that could be obtained from the lysate was estimated by adding saturating levels of GTP γ S (▼), which would activate most or all of the cdc42 in the cellular lysate. In this manner the levels of cdc42 in a cellular population can be quantified.

FIG. 17 provides photomicrographs of live cells injected with fluorescently labeled CBD. The warmer colors indicate where the activated
15 cdc42 is present within the cells. The merocyanin dye ("mero") is an environmentally sensitive dye that will fluoresce at higher intensity upon binding of the CBD-mero conjugate to activated cdc42. In this case, the intensity of fluorescence emitted by a non-environmentally sensitive fluorophore (Alexa) linked to CBD was determined relative to the intensity of fluorescence emitted
20 by the CBD-mero fluorophore. The ratio of the two permitted background fluorescence intensity to be subtracted so that the fluorescence from activated cdc-42 could be localized with greater precision.

Detailed Description

25 The following definitions are used, unless otherwise described.

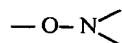
Alkylene, alkenylene, alkynylene, etc. denote both straight and branched groups; but reference to an individual radical such as "propylene" embraces only the straight chain radical; a branched chain isomer such as "isopropylene" being specifically referred to. Aryl denotes a phenyl radical or an ortho-fused bicyclic
30 carbocyclic radical having about nine to ten ring atoms in which at least one ring is aromatic.

The term "amino acid," includes the residues of the natural amino acids (e.g. Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Hyl, Hyp, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, and Val) in D or L form, as well as unnatural amino

acids (e.g. phosphoserine, phosphothreonine, phosphotyrosine, hydroxyproline, gamma-carboxyglutamate; hippuric acid, octahydroindole-2-carboxylic acid, statine, 1,2,3,4,-tetrahydroisoquinoline-3-carboxylic acid, penicillamine, ornithine, citruline, α -methylalanine, para-benzoylphenylalanine, phenylglycine, propargylglycine, sarcosine, and tert-butylglycine). The term also includes natural and unnatural amino acids bearing a conventional amino protecting group (e.g. acetyl or benzyloxycarbonyl), as well as natural and unnatural amino acids protected at the carboxy terminus (e.g. as a (C₁-C₆)alkyl, phenyl or benzyl ester or amide). Other suitable amino and carboxy protecting groups are known to those skilled in the art (See for example, T.W. Greene, *Protecting Groups In Organic Synthesis*; Wiley: New York, 1981, and references cited therein).

The term "peptide" includes any sequence of 2 or more amino acids. The sequence may be linear or cyclic. For example, a cyclic peptide can be prepared or may result from the formation of disulfide bridges between two cysteine residues in a sequence. Thus, the term includes proteins, enzymes, antibodies, oligopeptides, and polypeptides. Peptide sequences specifically recited herein are written with the amino terminus on the left and the carboxy terminus on the right.

An "aminoxy group" is a group having the following formula



wherein the open valences are filled by any acceptable radical. A "secondary aminoxy group" is an aminoxy group where one of the open valences on the nitrogen is filled by a radical other than a hydrogen.

The term "functional molecule" includes any compound that can be linked to a peptide to provide a peptide conjugate having useful properties. Such conjugates may be useful for studying the structure or function of the peptide, or a polypeptide, antibody, antigen or other protein. Functional groups linked to peptides by the present methods can be used in any assay, procedure or tracing protocol known to one of skill in the art. While the functional groups may be used as biosensors as contemplated below, their utility is not restricted to such use. Conjugates with such functional molecules may be useful for drug

screening, as pharmacological tools, as research tools, or as therapeutic agents. For example, the term functional molecule includes biophysical probes, peptides, polynucleotides, therapeutic agents, cross-linking groups (chemical or photochemical), a compound that modifies the biological activity of the peptide, or a caged molecule (e.g. a reporting molecule or a biologically active agent that is masked and that can be unmasked by photoactivation or chemical means).

The term "biophysical probe" includes any group that can be detected *in vitro* or *in vivo*, such as, for example, a fluorescent group, a phosphorescent group, a nucleic acid indicator, an ESR probe, another reporting group, a moiety, or a dye that is sensitive to pH change, ligand binding, or other environmental aspects.

Amino acids and peptides that include one or more aminooxy groups are useful intermediates for preparing peptide conjugates. The aminooxy group(s) can typically be positioned at any suitable position on the amino acid or peptide.

For example, the aminooxy group(s) can conveniently be incorporated into the side chain of the amino acid or into one or more side chains of the peptide.

Thus, as used herein with respect to the amino acids and peptides of the invention, the term "a radical comprising one or more aminooxy groups" includes any organic group that can be attached to the amino acid or peptide that includes one or more aminooxy groups. For example, the term includes a carbon chain having two to ten carbon atoms; which is optionally partially unsaturated (i.e. contains one or more double or triple bonds); which chain is optionally interrupted by one or more (e.g. 1, 2, or 3) -NH-, -O-, or -S-; which chain is optionally substituted on carbon with one or more (e.g. 1, 2, or 3) oxo (=O) groups; and which chain is optionally substituted with one or more (e.g. 1, 2, or 3) aminooxy groups. Preferably, the aminooxy group(s) are secondary aminooxy groups.

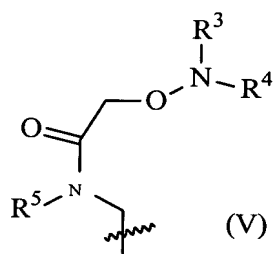
The term "cross-linking group" refers to any functionality that can form a bond with another functionality, such as photoaffinity label or a chemical crosslinking agent.

The term "caged molecule" includes a molecule or reporter group that is masked such that it can be activated (i.e. unmasked) at a given time or location of choice, for example using light or a chemical agent.

Specific and preferred values listed below for radicals, substituents, and ranges, are for illustration only; they do not exclude other defined values or other values within defined ranges for the radicals and substituents.

Specifically, (C₁-C₆)alkylene can be methylene, ethylene, propylene, isopropylene, butylene, iso-butylene, sec-butylene, pentylene, or hexylene; (C₃-C₈)cycloalkyl can be cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, or cyclooctyl; (C₂-C₆)alkenylene can be vinylene, allylene, 1-propenylene, 2-propenylene, 1-butenylene, 2-butenylene, 3-butenylene, 1-pentenylene, 2-pentenylene, 3-pentenylene, 4-pentenylene, 1-hexenylene, 2-hexenylene, 3-hexenylene, 4-hexenylene, or 5-hexenylene; (C₂-C₆)alkynylene can be ethynylene, 1-propynylene, 2-propynylene, 1-butyne, 2-butyne, 3-butyne, 1-pentyne, 2-pentyne, 3-pentyne, 4-pentyne, 1-hexynylene, 2-hexynylene, 3-hexynylene, 4-hexynylene, or 5-hexynylene; and aryl can be phenyl, indenyl, or naphthyl;

A specific value for R is a radical of formula (V):



wherein

R³ is hydrogen, (C₁-C₆)alkyl, an amino protecting group, or a radical comprising one or more aminoxy groups;

R⁴ is hydrogen, or an amino protecting group; and

R⁵ is hydrogen, or (C₁-C₆)alkyl.

A specific value for R¹ is hydrogen or benzyloxycarbonyl.

A specific value for R² is hydrogen.

A specific value for R³ is methyl.

A specific value for R⁴ is hydrogen, 2-chlorobenzyloxycarbonyl, or benzyloxycarbonyl.

A specific value for R⁵ is hydrogen.

A specific value for R⁶ is an antibody.

A specific value for R^6 is a peptide or polypeptide or antibody that includes about 2 to about 1000 amino acids. A more specific value for R^6 is a peptide that includes about 5 to about 500 amino acids. An even more specific value for R^6 is a peptide that includes about 10 to about 100 amino acids.

5 Specifically X is a linking group that is about 5 angstroms to about 100 angstroms in length. More specifically, X is a linking group of about 5 angstroms to about 25 angstroms in length.

Specifically X is $-R_a-C(=O)-NH-R_b-$ wherein each of R_a and R_b is independently (C_1-C_6) alkylene. Preferably, each of R_a and R_b is methylene (-
10 CH_2-).

A preferred value for R^6 is KKKEKERPEISLPSDFEHTIHVGF
DACTGEFTGMPEQWARLLQT (SEQ ID NO: 1) or an antibody.

A specific value for R^7 is hydrogen.

Another specific value for R^7 is (C_1-C_6) alkyl.

15 A preferred value for R^7 is methyl.

Specifically Y is a linking group that is about 5 angstroms to about 100 angstroms in length. More specifically, Y is a linking group of about 5 angstroms to about 25 angstroms in length.

A specific value for Y is (C_1-C_6) alkylene.

20 A preferred value for Y is methylene ($-CH_2-$).

Fluorescent Dyes

Any fluorescent dye known to one of skill in the art is contemplated by the present invention as a functional group. A fluorescent dye can be excited to
25 fluoresce by exposure to a certain wavelength of light. When used as the reporter molecule of the biosensors of the present invention, the dye is preferably environmentally sensitive. As used herein, "environmentally sensitive" means that the signal from the functional group changes when the peptide, polypeptide or antibody interacts with, or becomes exposed to, a different environment. For
30 example, when the environmentally sensitive functional group is a fluorescent dye, the fluorescence from that fluorescent dye will change as the environment changes. In one embodiment an environmentally sensitive fluorescent dye attached to a peptide, polypeptide or antibody will fluoresce differently upon target binding by the peptide, polypeptide or antibody to which the dye is

attached. Any dye which emits fluorescence and whose fluorescence changes when the pH or the hydrophilicity/hydrophobicity of the environment changes is an environmentally sensitive dye contemplated by the present invention.

Preferred fluorescent groups include molecules that are capable of
5 absorbing radiation at one wavelength and emitting radiation at a longer wavelength, such as, for example, Alexa-532, Hydroxycoumarin, Aminocoumarin, Methoxycoumarin, Coumarin, Cascade Blue, Lucifer Yellow, P-Phycoerythrin, R-Phycoerythrin, (PE), PE-Cy5 conjugates, PE-Cy7 conjugates, Red 613, Fluorescein, BODIPY-FL, BODIPY TR, BODIPY TMR,
10 Cy3, TRITC, X-Rhodamine, Lissamine Rhodamine B, PerCP, Texas Red, Cy5, Cy7, Allophycocyanin (APC), TruRed, APC-Cy7 conjugates, Oregon Green, Tetramethylrhodamine, Dansyl, Dansyl aziridine, Indo-1, Fura-2, FM 1-43, DilC18(3), Carboxy-SNARF-1, NBD, Indo-1, Fluo-3, DCFH, DHR, SNARF, Monochlorobimane, Calcein, N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl) amine
15 (NBD), ananilinonaphthale, deproxyl, phthalamide, amino pH phthalamide, dimethylamino-naphthalenesulfonamide, probes comparable to Prodan, Lordan or Acrylodan and derivatives thereof. Coumarin fluorescent dyes include, for example, amino methylcoumarin, 7-diethylamino-3-(4'-(1-maleimidyl)phenyl)-4-methylcoumarin (CPM) and N-(2-(1-maleimidyl)ethyl)7-
20 diethylaminocoumarin-3-carboxamide (MDCC). Preferred fluorescent probes are sensitive to the polarity of the local environment and are available to those of skill in the art.

Other useful functional groups include those that display fluorescence resonance energy transfer (FRET). Many such donor-acceptor pairs are known,
25 and include fluorescein to rhodamine, coumarin to fluorescein or rhodamine, etc. Still another class of useful label pairs include fluorophore-quencher pairs in which the second group is a quencher, which decreases the fluorescence intensity of the fluorescent group. Some known quenchers include acrylamide groups, heavy atoms such as iodide and bromate, nitroxide spin labels such as TEMPO,
30 etc. These can be adapted for use as environmentally sensitive functional groups of biosensors.

Exemplary fluorescent proteins which can be used to label the present peptides, polypeptides and antibodies include green fluorescent protein (GFP),

cyan fluorescent protein (CFP), red fluorescent protein (RFP), yellow fluorescent protein (YFP), enhanced GFP (EGFP), enhanced YFP (EYFP), and the like.

New Fluorescent Dyes

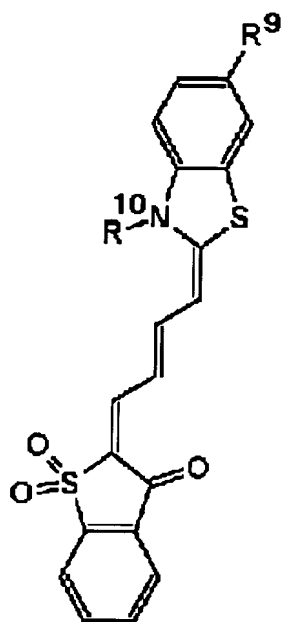
- 5 The present invention also provides novel fluorescent dyes that retain high fluorescence emission after conjugation to other molecules and avoid problems of aggregation and insolubility. These dyes are particularly preferred for many of the imaging methods and conjugates contemplated but need not be restricted to use in the methods and conjugates contemplated herein. Thus, the
- 10 present invention is directed to highly fluorescent dyes of structure IV wherein,
- each m is separately an integer ranging from 1-3;
n is an integer ranging from 0 to 5;
R⁸, R¹¹ and R¹² are separately CO, SO₂, C=C(CN)₂, S, O or C(CH₃)₂;
- 15 each R¹³ is alkyl, branched alkyl or heterocyclic ring derivatized with charged groups to enhance water solubility and enhance photostability;
- R⁹ and R¹⁰ are chains carrying charged groups to enhance water solubility (i.e. sulfonate, amide, ether) and/or chains bearing reactive groups for conjugation to other molecules. The reactive group is a functional group that is
- 20 chemically reactive (or that can be made chemically reactive) with functional groups typically found in biological materials, or functional groups that can be readily converted to chemically reactive derivatives using methods well known in the art. In one embodiment of the invention, the charged and reactive groups are separately haloacetamide (-NH-(C=O)-CH₂-X), where X is Cl, Br or I.
- 25 Alternatively, the charged and reactive groups are separately amine, maleimide, isocyanato (-N=C=O), isothiocyanato (-N=C=S), acyl halide, succinimidyl ester, or sulfosuccinimidyl ester. In another embodiment, the charged and reactive groups are carboxylic acid (COOH), or derivatives of a carboxylic acid. An appropriate derivative of a carboxylic acid includes an alkali or alkaline earth
- 30 metal salt of carboxylic acid. Alternatively, the charged and reactive groups are reactive derivatives of a carboxylic acid (-COOR_x), where the reactive group R_x is one that activates the carbonyl group of -COOR_x toward nucleophilic displacement. In particular, R_x is any group that activates the carbonyl towards nucleophilic displacement without being incorporated into the final displacement

product. Examples of COORx : ester of phenol or naphtol that is further substituted by at least one strong electron withdrawing group, or carboxylic acid activated by carbodiimide, or acyl chloride, or succinimidyl or sulfosuccinimidyl ester. Additional charged and reactive groups include, among others, sulfonyl
5 halides, sulfonyl azides, alcohols, thiols, semicarbazides, hydrazines or hydroxylamines.

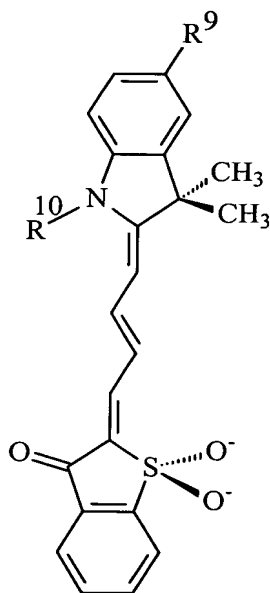
For all dyes of the invention, any net positive or negative charges possessed by the dye are balanced by a biologically compatible counterion or counterions. As used herein, a substance that is biologically compatible is not
10 toxic as used, and does not have a substantially deleterious effect on biomolecules. Examples of useful counterions for dyes having a net negative charge include, but are not limited to, alkali metal ions alkaline earth metal ions, transition metal ions, ammonium and substituted ammonium ions. Examples of useful counterions for dyes having a net positive charge include, but are not
15 limited to chloride, bromide, iodide, sulfate, phosphate, perchlorate, nitrate, tetrafluoroborate.

As used herein, R⁹ and R¹⁰ chains for conjugation are alkyl, of unlimited length, preferably with 1-6 carbons, and can include other moieties such as ether, amide, or sulfonate to improve water solubility. The groups can be substituted
20 on the end or in the middle of the alkyl chain.

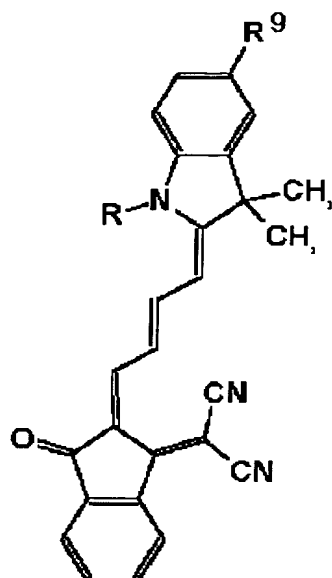
In one embodiment, the fluorescent dye of the present invention has the following structure (VI):



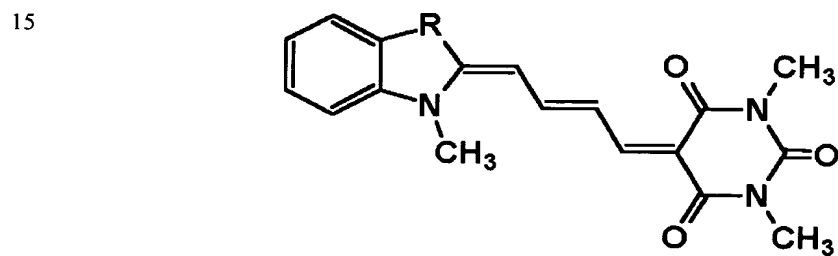
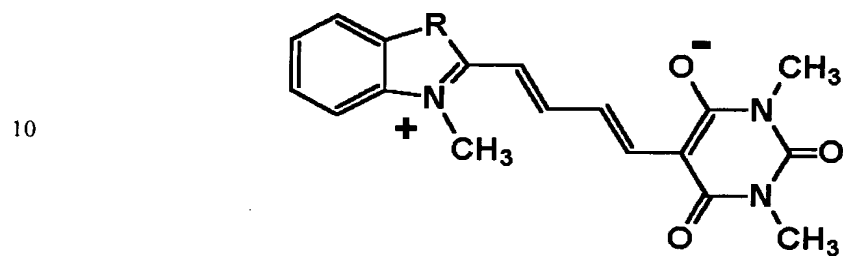
In another embodiment, the fluorescent dye of the present invention has
 5 the following structure (VII):



In another embodiment, the fluorescent dye of the present invention has
 the following structure (VIII):



Depending upon their environment, the fluorescent dyes of the present invention can exist in somewhat different polarization states. This property can modulate the solubility and the emission wavelength of the dye. For example,
 5 the fluorescent dye depicted below can be charged or non-charged.





The fluorescent dyes of the present invention can absorb and emit light at a variety of wavelengths, depending on the arrangement and variety of substituents employed. Thus, the choice of whether to use R^8 , R^{11} and R^{12} as $-\text{CO}-$, $-\text{SO}_2-$, $-\text{C}=\text{C}(\text{CN})_2-$, $-\text{S}-$, $-\text{O}-$ or $-\text{C}(\text{CH}_3)_2-$ can influence the absorption and emission wavelength. However, by varying the substituents of the present invention, one of skill in the art can readily ascertain which combination of substituents will yield a fluorescent dye with a desired absorption and emission spectrum.

Moreover, according to the present invention, the degree of conjugation of the dye, and in particular, the length of the alkylene chain connecting the two rings, can predictably influence the absorption and emission wavelength of the dye. Thus, addition of one $-\text{C}=\text{C}-$ group can shift the fluorescence wavelength about +100 nm. Smaller incremental changes in the emission wavelength can be made by adding a conjugated group to one of the rings in the dye. Thus, one of skill in the art can readily modulate the emission wavelength of the dye as desired. In one embodiment, the absorption and emission wavelengths can be altered to range from about 300 nm to about 800 nm. Preferably, the absorption and emission wavelengths of the present dyes range from about 450 nm to higher wavelengths. Any variety of methods can be used to make the present dyes.

Preferred nucleic acid indicators include intercalating agents and oligonucleotide strands, such as, for example, YOYO-1, Propidium Iodide, Hoechst 33342, DAPI, Hoechst 33258, SYTOX Blue, Chromomycin A3, Mithramycin, SYTOX Green, SYTOX Orange, Ethidium Bromide, 7-AAD, Acridine Orange, TOTO-1, TO-PRO-1, Thiazole Orange, Propidium Iodide, TOTO-3, TO-PRO-3, LDS 751.

25

Synthon and Peptide Intermediates

The synthetic intermediates (i.e. synthons) of the invention that include one or more aminoxy groups can be incorporated into peptides using a variety of techniques that are known in the art. For example, as discussed below, the synthons can be incorporated into a peptide using solid-phase peptide synthesis, solution-phase peptide synthesis, native chemical ligation, intein-mediated protein ligation, and chemical ligation.

Peptides may be prepared using solid-phase peptide synthesis (SPPS). For example, according to the SPPS technique, protected amino acids in organic

solvents can be added one at a time to a resin-bound peptide chain, resulting in the assembly of a target peptide having a specific sequence in fully-protected, resin-bound form. The product peptide can then be released by deprotection and cleavage from the resin support (Wade, L.G., JR., Organic Chemistry 4th Ed.

5 (1999)). As illustrated in Example 2 below, amino acids containing an aminooxy functional group can be incorporated into peptides using SPPS. Use of this methodology allows an amino acid containing an aminooxy functional group to be positioned at a desired location within a synthesized peptide chain.

Amino acids containing an aminooxy group can also be incorporated into
10 a peptide using solution-phase peptide synthesis (Wade, L.G., JR. Organic Chemistry 4th Ed. (1999)). Solution-phase peptide synthesis involves protecting the amino-terminus of a peptide chain followed by activation of the carboxyl-terminus allowing the addition of an amino acid or a peptide chain to the carboxy-terminus (Wade, L.G., JR. Organic Chemistry 4th Ed. (1999)).

15 Native chemical ligation is a procedure that can be used to join two peptides or polypeptides together thereby producing a single peptide or polypeptide having a native backbone structure. Native chemical ligation is typically carried out by mixing a first peptide with a carboxy-terminal α -thioester and a second polypeptide with an amino-terminal cysteine (Dawson,
20 P.E., et al., (1994), Science 266:776-779; Cotton, G.J., et al., (1999), J. Am. Chem. Soc. 121:1100-1101). The thioester of the first peptide undergoes nucleophilic attack by the side chain of the cysteine residue at the amino terminus of the second peptide. The initial thioester ligation product then undergoes a rapid intramolecular reaction because of the favorable geometric
25 arrangement of the alpha-amino group of the second peptide. This yields a product with a native peptide bond at the ligation site. A polypeptide beginning with cysteine can be chemically synthesized or generated by intein vectors, proteolysis, or cellular processing of the initiating methionine. This method allows mixing and matching of chemically synthesized polypeptide segments.

30 The synthons of the invention are particularly useful in combination with native chemical ligation, because native chemical ligation allows a synthetic peptide having a specifically positioned amino acid (e.g. a synthon of the invention) to be selectively ligated to other peptides or into a larger polypeptide, antibody or protein. The ability to specifically incorporate aminooxy modified

amino acids into a peptide chain allows useful moieties to be linked at any position within a peptide, polypeptide, antibody or protein. Examples of such moieties that can be incorporated into a peptide using this method include, but are not limited to, phosphorylated or glycosylated amino acids, unnatural amino acids, tags, labels, crosslinking reagents, biosensors, reactive groups, and fluorophores. Another advantage of native chemical ligation is that it allows incorporation of peptides into a polypeptide that are unable to be added by ribosomal biosynthesis.

Intein-mediated protein ligation may also be used to selectively place amino acids containing aminooxy functional groups into peptides. Inteins are intervening sequences that are excised from precursor proteins by a self-catalytic mechanism and thereby expose reactive ends of a peptide. Intein vectors have been developed that not only allow single-step purification of proteins, but also yield polypeptides with reactive ends necessary for intein-mediated protein ligation (IPL) (also called expressed protein ligation)(EPL) (Perler, F.R. and Adam, E., (2000) *Curr. Opin. Biotechnol.* 11(4):377-83; and Evans, T.C., et al., (1998) *Protein Sci* 7:2256-2264). This method allows a peptide having a selectively placed amino acid containing an aminooxy functional group to be readily ligated to any peptide with reactive ends generated by intein excision.

Two peptides or polypeptides may also be linked through use of chemical ligation. Chemical ligation occurs when two peptide segments are each linked to functional groups that react with each other to form a covalent bond producing a non-peptide bond at the ligation site (Wilken, J. and Kent, S.B.H., (1998) *Curr. Opin. Biotechnol.* 9:412-426). This method can be used to ligate a peptide having a specifically positioned aminooxy functional group to another peptide or polypeptide to produce a desired polypeptide that may be later linked to a detectable group.

A functional molecule ("D") can be attached to a peptide comprising an aminooxy group through a direct linkage (e.g. an amide bond -O-N-C(=O)-D) or through a linking group. The structure of the linking group is not crucial, provided it does not interfere with the use of the resulting labeled peptide. Preferred linking groups include linkers that separate the aminooxy nitrogen and the detectable group by about 5 angstroms to about 100 angstroms. Other

preferred linking groups separate the aminooxy nitrogen and the detectable group by about 5 angstroms to about 25 angstroms.

For example, the linking group can conveniently be linked to the detectable group through an: 1) amide (-N(H)C(=O)-, -C(=O)N(H)-), 2) ester (-OC(=O)-, -C(=O)O-), 3) ether (-O-), 4) thioether (-S-), 5) sulfinyl (-S(O)-), or 6) sulfonyl (-S(O)₂) linkage. Such a linkage can be formed from suitably functionalised starting materials using synthetic procedures that are known in the art.

The linking group can conveniently be linked to the nitrogen of the aminooxy group to form an amide (-O-N(H)C(=O)- or a thiourea linkage (-O-N-C(=S)-N-) linkage, using reagents and conditions that are known in the art.

The aminooxy group can be attached to a peptide through a direct bond (e.g. a carbon-oxygen bond) between the aminooxy oxygen and a side chain of the peptide, or the aminooxy group can be attached to the peptide through a linking group. The structure of the linking group is not crucial, provided it does not interfere with the use of the resulting labeled peptide. Preferred linking groups include linking groups that separate the aminooxy oxygen and the side chain of the peptide by about 5 angstroms to about 100 angstroms. Other preferred linking groups separate the aminooxy oxygen and the side chain of the peptide by about 5 angstroms to about 25 angstroms.

A specific linking group (e.g. X or Y) can be a divalent (C₁-C₆)alkylene, (C₂-C₆)alkenylene, or (C₂-C₆)alkynylene chain, or a divalent (C₃-C₈)cycloalkyl, or aryl ring.

Thus, a simple and efficient synthetic methodology for site-specific labeling of peptides after synthesis has been developed that provides high yield, selectivity and compatibility with both solid-phase peptide synthesis and C^α-thioester peptides. The approach and primary advantages can be summarized as follows:

(1) A protected aminooxy amino acid that can be incorporated into peptides has been synthesized;

(2) Procedures have been optimized to yield highly efficient and specific modification of the aminooxy nitrogen in the presence of unprotected competing nucleophiles, including cysteine, lysine and amino groups;

(3) One preferred electrophile that can be used for labeling, an activated carboxylic ester, is readily available in the majority of commercially available fluorescent dyes and labels;

(4) Labeling of the aminooxy group occurs after synthesis and
5 purification, thus enabling the use of chemically sensitive fluorophores and labels that would otherwise not survive earlier synthetic procedures;

(5) The synthetic methodology is compatible with the steps required for the synthesis of proteins by total chemical synthesis or expressed protein ligation, namely synthesis of C^α-peptide thioesters and amide-forming ligations;
10 and

(6) Combinatorial screening of both the functional molecule and its placement will enable the rapid synthesis of optimally labeled polypeptide-based biosensors.

15 **Biosensors**

Using the methods of the present invention with the synthons and peptides provided herein, antibodies, antigens and other polypeptides can be labeled with or attached to functional groups. Such labeled antibodies and polypeptides have particular utility as biosensors. As used herein, a “biosensor”
20 is a peptide, antigen, polypeptide or antibody with an environmentally sensitive functional group. For example, the fluorescence emitted by an environmentally-sensitive fluorescent dye can change when the protein biosensor becomes exposed to a different environment. Such a change in environment can occur, for example, when the biosensor binds to, or associates with, a target protein or
25 cellular structure. The change in fluorescence can be used to quantify or otherwise monitor the amount of binding or interaction between the biosensor and the target site. The Examples provided herein further illustrate how such a biosensor can be used.

The present invention provides methods for identifying an optimal
30 position for the functional group on the peptide, polypeptide or antibody which involve generating a series of peptides, each peptide have the same amino acid sequence and functional group. However, the functional group is positioned at a different location along the backbone of each peptide in the series. To determine which location is best, the strength of the signal from the various peptides is

observed under different conditions and the optimal location provides optimal functioning by the functional group and the peptide, polypeptide or antibody. For example, when the functional group provides a signal, a stronger signal is preferred so long as the function and the chemical and physical properties of the peptide, protein or antibody are not impaired. When an environmentally sensitive functional group is chosen, a maximal change in signal is preferred as the environment is changed. For example, when the environmentally sensitive functional group is attached to detect an interaction of the peptide with a target, a maximal change in signal is preferred when the peptide binds or interacts with its target, unless of the location of the functional group affects the binding affinity, binding selectivity or another desirable attribute of the peptide.

Thus to determine an optimal location for a functional group in an antibody or polypeptide which can bind to a target, a series of peptides are first synthesized, each with the functional group at a different position. The peptides can be incorporated into the polypeptide or antibody using the methods described herein to generate a series of biosensors, each with a functional group at a somewhat different position. The interaction of the different polypeptide or antibody biosensors with target is observed. In general, an optimal position for the functional group on such a biosensor is that position which permits stable and selective binding to target with a maximal signal change upon binding.

However, some variability in binding and signal strength can be tolerated so long as an observable, localized signal is detected upon interaction of the biosensor and the target. Thus, the strength of binding between the biosensor and target should be sufficient to permit observation of bound biosensor. If the biosensor is only transiently bound, little or no localized signal from the probe may be observed; instead, only diffuse signal from biosensor in solution may be observed. Similarly, if the biosensor is bound non-selectively, non-localized signal may be detected from many sites. Obviously, a strongly localized signal that clearly correlates with biosensor binding to target is preferred. A readily detectable change in signal strength or quality as the biosensor and target interact is also preferred.

Procedures known to one of skill in the art can be used to detect a signal provided by the functional group and to correlate a change in signal by an environmentally sensitive functional group. Signals contemplated by the

present invention include fluorescent emissions, radioactive emissions, enzymatic production of a colored product, and the like. One of skill in the art can readily detect these signals using a fluorescence microscope, a scintillation counter, a light or radioactively sensitive photoemulsion, a light microscope, a spectrophotometer or other means. When an environmentally-sensitive functional group is used, the signal can change in lifetime, strength, color or other quality to signal interaction of the functional group with the environment. For example, when the environmentally-sensitive functional group is a fluorescent dye, the signal change can be a color, wavelength, intensity or lifetime of fluorescence emitted by the dye. The change need only be detectable, for example, a change in wavelength of fluorescence of about 50 nm to about 300 nm can be readily detected. However, a change in wavelength of greater than about 10 nm is preferred. More preferably the change in wavelength is greater than 20 nm. In one embodiment the change in wavelength can vary from about 30 nm to about 200 nm.

Where convenient, a peptide, as opposed to a polypeptide or antibody, with an attached functional group may be a biosensor, for example, when the binding affinity and selectivity of the peptide is representative of the larger protein of which it is normally a part. Alternatively, the labeled peptide can be ligated into a polypeptide to form the protein or antibody of which it is normally a part using the methods described herein. In one embodiment, the peptide is incorporated onto one of the termini of a protein, for example, through procedures described in Curr. Op. Biotechnology 9:412 (1998) or Ann. Rev. Biochem. 57:957-89 (1988). In another embodiment, the peptide can be incorporated into the middle of a protein, for example, by using the procedures described in PNAS 95:6705 (1998).

Targets

As used herein, a "target" is any molecule, structure or complex that a peptide, polypeptide or antibody linked to a functional group by the present methods can interact with. Targets contemplated by the present invention include antigens, antibodies, proteins, enzymes, membrane proteins, structural proteins, major histocompatibility proteins, DNA binding proteins, receptors, ligands, cofactors, nucleic acids, kinases, GTPases, ATPases, proteins involved

09839577, 042001

in motility and the like. Targets may undergo structural changes and other types of changes, for example, phosphorylation, de-phosphorylation, conformational changes, ligand binding, co-factor binding, activation, post-translational modification, carbohydrate or sugar attachment, membrane interactions and the like. Specific targets include calmodulin, Rho GTPases, rac, cdc42, actin, myosin, major histocompatibility proteins.

Antigen and Antibody Targets

When antibodies are linked to functional groups using methods of the present invention the target can be an antigen. Alternatively, an antigen or a peptide epitope can serve as a biosensor for an antibody target. The present invention contemplates use or detection of any antigen or antibody known to one of skill in the art as a target.

In general, a molecule must have sufficient complexity and sufficient molecular weight in order to act as an antigen. In order to have sufficient complexity, the antigen must have at least one epitope. As used in this invention, the term "epitope" is meant to include any determinant capable of specific interaction with the monoclonal antibodies of the invention. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics.

In order to have a sufficient molecular weight, the antigen generally must have a molecular weight that is greater than 2,000 daltons. Formerly, it was thought that the lower molecular weight limit to confer antigenicity was about 5,000 daltons. However, antigenicity has recently been demonstrated with molecules having molecular weights as low as 2,000 daltons. Molecular weights of 3,000 daltons and more appear to be more realistic as a lower limit for immunogenicity, and approximately 6,000 daltons or more is preferred.

In preparing antigens to produce antibodies for attachment to functional group, it is desirable to use antigens with a high degree of purity. Accordingly, it is desirable to use a purification process permitting isolation of the antigen from antigenically distinct materials. Antigenically distinct materials are undesired large molecules that may compete with the target antigen for antibody production thereby minimizing production of the desired antibodies or inducing

cross-reactive antibodies of low specificity or affinity. The practice of the invention can accordingly include a number of purification steps using available techniques. Purification can, for example, be effected by size exclusion chromatography, ion exchange chromatography, dialysis, cold organic solvent extraction, gel electrophoresis and/or fractional crystallization means which are available to one of skill in the art. Preferably an antigen used for antibody preparation is at least about 90% pure and more preferably at least about 99% pure.

Removal of small molecule reactants and reaction products from the synthesized antigen is generally desirable. However, some small molecular substances may be useful, for example, for control of pH and salinity. Thus, a convenient end-product form in which to recover the antigen is, in a buffered aqueous solution that is suitable for direct administration to animals.

Antibodies

The present invention contemplates linkage of any available antibody to functional groups. Such antibodies can be polyclonal or monoclonal antibodies. The term "antibody" as used in this invention is meant to include intact antibody molecules as well as fragments thereof, such as, for example, Fab and F(ab')₂, which are capable of binding to the antigen or its epitopic determinant.

Polyclonal antibodies can be raised by administration of an antigen of the invention to vertebrate animals, especially mammals such as goats, rabbits, rats or mice using known immunization procedures. Usually a buffered solution of the antigen accompanied by Freund's adjuvant is injected subcutaneously at multiple sites. A number of such administrations at intervals of days or weeks is usually necessary. A number of animals, for example from 3 to 20, are injected with the expectation that only a small proportion will produce desirable antibodies. However, one animal can provide antibodies sufficient for thousands of assays. Antibodies are recovered from animals after some weeks or months.

Exemplary immunogenic carrier materials can be used with the antigen to enhance the immune response. The carrier material can be a natural or synthetic substance, provided that it is an antigen or a partial antigen. For example, the carrier material can be a protein, a glycoprotein, a nucleoprotein, a polypeptide, a polysaccharide, a lipopolysaccharide, or a polyamino acid. *See also*, the carrier

molecules and antibody production methods set forth in Cremer et al., "Methods in Immunology" (1963), W. A. Benjamin Inc., New York, pp. 65-113 and Harlow and Lane, "Antibodies: A Laboratory Manual" (1988), Cold Spring Harbor Laboratory, New York, p.5. These disclosures are herein incorporated by
5 reference.

A preferred class of natural carrier materials is the proteins. Proteins can be expected to have a molecular weight in excess of 5,000 daltons, commonly in the range of from 34,000 to 5,000,000 daltons. Specific examples of such natural proteins are hen ovalbumin (OA), bovine serum albumin (BSA), keyhole limpet
10 hemocyanin (KLH), horse gammaglobulin (HGG), and thyroglobin.

Exemplary of synthetic carrier is the polyamino acid, polylysine. Where the synthetic antigen comprises a partially antigenic carrier conjugated with a hapten, it will generally be desirable for the conjugate to have a molecular weight in excess of 6,000 daltons, although somewhat lower molecular weights
15 may be useful. Preferably, the natural carrier has some solubility in water or aqueous alcohol. Desirably, the carriers are nontoxic to the animals to be used for generating antibodies.

The carrier can be coupled to antigens by any available means, included the procedures provided herein. Preferably, a carrier moiety has a plurality of
20 hapten or antigen moieties coupled to it, for example, about 15 to 30 for a protein of 100,000 daltons. While steric hindrance and reduced structural complexity may reduce the number of haptens or antigens attached to the carrier, the maximum number is preferred. For example, up to about 25 to about 50 hapten moieties can be coupled to BSA carriers.

25 The use of monoclonal antibodies as biosensors of this invention is preferred because monoclonal antibodies are homogeneous and can be continuously produced in large quantities. Monoclonal antibodies are prepared by recovering lymph node or spleen cells from immunized animals and immortalizing the cells in conventional fashion, e.g., by fusion with myeloma
30 cells or by Epstein-Barr virus transformation. Clones expressing the desired antibody are identified by screening cell line media for reactivity with the antigen used to immunize the animals. One of skill in the art can use readily available methods to make monoclonal antibodies, for example, by using the hybridoma technique described originally by Koehler and Milstein, Eur. J.

Immunol., 6:511 (1976), by Hammerling et al., in "Monoclonal Antibodies and T-Cell Hybridomas", Elsevier, New York, pp. 563-681 (1981), and by Zola, in "Monoclonal Antibodies: A Manual of Techniques", CRC Press, Boca Raton, Fla. (1987). The hybrid cell lines can be maintained in vitro in cell culture media. The cell lines producing the antibodies by these procedures can be selected and/or maintained in a medium containing hypoxanthine-aminopterin thymidine (HAT). However, once the hybridoma cell line is established, it can be maintained on a variety of nutritionally adequate media. Moreover, the hybrid cells lines can be stored and preserved in any number of conventional ways, including freezing and storage under liquid nitrogen. Frozen cell lines can be revived and cultured indefinitely with resumed synthesis and secretion of monoclonal antibody.

The secreted antibody is recovered from tissue culture supernatant or ascites fluid by conventional methods such as immune precipitation, ion exchange chromatography, affinity chromatography such as protein A/protein G column chromatography, or the like. The antibodies described herein are also recovered from hybridoma cell cultures by conventional methods such as precipitation with 50% ammonium sulfate. If desired, the purified antibodies can then be sterile filtered before use.

The term "monoclonal antibody" as used herein refers to any antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations, which typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they are synthesized by the hybridoma culture, uncontaminated by other immunoglobulins.

The monoclonal antibodies herein also include hybrid and recombinant antibodies produced by splicing a variable (including hypervariable) domain of an anti-adduct antibody with a constant domain (e.g. "humanized" antibodies), or a light chain with a heavy chain, or a chain from one species with a chain from

another species, or fusions with heterologous proteins, regardless of species of origin or immunoglobulin class or subclass designation, as well as antibody fragments (e.g., Fab, F(ab')₂ and Fv), so long as they exhibit the desired biological activity. See e.g. Cabilly et al. U.S. Pat. No. 4,816,567; Mage and
5 Lamoyi, "Monoclonal Antibody Production Technique and Applications", pp. 79-97 (Marcel Dekker, Inc., New York, 1987).

Thus, the modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogenous population of antibodies, and is not to be construed as requiring production of the antibody by any particular
10 method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method described by Koehler and Milstein, supra, or may be made by recombinant DNA methods (Cabilly, et al. supra).

In one embodiment, the biosensor antibodies or the present invention are
15 used for diagnostic or imaging purposes in vivo, within a mammalian subject. While the in vivo use of a monoclonal antibody from a foreign donor species in a different host recipient species is usually uncomplicated, an adverse immunological response by the host to antigenic determinants present on the donor antibody can sometimes arise. In some instances, this adverse response
20 can be so severe as to curtail the in vivo use of the donor antibody in the host. Further, the adverse host response may serve to hinder the intercellular adhesion-suppressing efficacy of the donor antibody. Methods to avoid such adverse reactions are available. For example, humanized antibodies or chimeric antibodies (Sun, et al., Hybridoma, 5 (Supplement 1):S17, 1986; Oi, et al., Bio
25 Techniques, 4(3):214, 1986) can be used. Chimeric antibodies are antibodies in which the various domains of the antibodies' heavy and light chains are coded for by DNA from more than one species. Typically, a chimeric antibody will comprise the variable domains of the heavy (V_H) and light (V_L) chains derived from the donor species producing the antibody of desired antigenic specificity,
30 and the constant domains of the heavy (C_H) and light (C_L) chains derived from the host recipient species. It is believed that by reducing the exposure of the host immune system to the antigenic determinants of the donor antibody domains, especially those in the C_H region, the possibility of an adverse immunological response occurring in the recipient species will be reduced. Thus, for example, it

is possible to produce a chimeric antibody for in vivo clinical use in humans which comprises mouse V_H and light V_L domains coded for by DNA isolated from ATCC HB X, and C_H and C_L domains coded for with DNA isolated from a human leukocyte.

5 The present invention further provides a composition comprising an antibody with a functional group attached by the methods of the present invention and a suitable carrier. Further, the present invention also provides a therapeutic composition comprising an effective amount of the antibody-functional group conjugate produced by the present methods and a
10 pharmaceutically acceptable carrier.

As used herein, the term "pharmaceutically acceptable carrier" encompasses any of the standard pharmaceutically accepted carriers, such as phosphate buffered saline solution, water, emulsions such as an oil/water emulsion or a triglyceride emulsion, various types of wetting agents, tablets,
15 coated tablets and capsules. An example of an acceptable triglyceride emulsion useful in intravenous and intraperitoneal administration of the compounds is the triglyceride emulsion commercially known as Intralipid RTM. Typically such carriers contain excipients such as starch, milk, sugar, certain types of clay, gelatin, stearic acid, talc, vegetable fats or oils, gums, glycols, or other known
20 excipients. Such carriers may also include flavor and color additives or other ingredients.

The invention will now be illustrated by the following non-limiting Examples.

25 **Example 1. Site-specific labeling of a secondary aminooxy group.**

Under controlled pH conditions, the low pK_a and enhanced nucleophilicity of an aminooxy group relative to other nucleophilic side chains found in peptides suggested the possibility of site-specific reaction with standard electrophiles such as succinimidyl esters (Figure 1). While selective labeling of a
30 primary aminooxy group in the context of an unprotected peptide was achieved, extensive attempts to utilize the primary aminooxy group during synthesis failed. Even when protected as the 2-chlorobenzyloxycarbonyl carbamate, deprotonation of the primary aminooxy group allowed rapid acylation, so it

0983957, 042001

could not be readily incorporated during peptide synthesis. Thus, under certain conditions, the use of a secondary aminooxy group may be preferred.

A test peptide containing both a secondary aminooxy group and nucleophilic amino acids that were most likely to interfere with selective labeling at the aminooxy nitrogen (lysine, cysteine, and the amino-terminus) was prepared. As illustrated in Figure 2, NH₂-AKAARAAAAK*AARACA-CO₂H (SEQ ID NO: 2), here designated SA-test peptide, was synthesized by incorporation and deprotection of N-(2-Cl-benzyloxycarbonyl)-N-methylaminooxy acetic acid (Figure 4, 3) during solid phase peptide synthesis.

The reactivity of the protected secondary aminooxy group was sufficiently attenuated to remain unreactive during Boc solid-phase peptide synthesis on thioester-linker resins. The 2-Cl-Z protection for the N-methylaminooxy amino acid was efficiently removed by standard HF cleavage procedures.

Conditions for selectively labeling the secondary aminooxy group were determined by varying the reaction pH and dye stoichiometry. Labeling with the succinimide ester of tetramethylrhodamine (TMR-OSu) was determined to be optimal in a solvent system consisting of 50%DMSO/50% aqueous acetate buffer at pH of 4.7 with 2 equivalents of dye per mole of peptide (Table 1). The crude reaction products were separated from unreacted dye, and characterized by RP-HPLC and ESI-MS. Under these conditions, a single molecule of dye was incorporated on the aminooxy group with a 78% yield, based on HPLC quantification. However, side reaction products were also isolated and determined to be either SA-test peptide labeled with two dye molecules (~13%) or acetylated peptide products (~5%). The selectivity, as defined by the ratio of the peak areas of desired single-labeled product over double-labeled products, was 6/1 (Table 1).

Table 1: Labeling of SA-Test Peptide

<u>Buffers</u>				<u>Dye</u>	<u>Percentage</u>			
<u>Reaction</u>	<u>Gel</u>	<u>TCEP</u>	<u>pH*</u>	<u>Equivale</u>	<u>S</u>	<u>1Dy</u>	<u>2Dy</u>	<u>Selectivi</u>
	<u>Filtration</u>			<u>nts</u>	<u>M</u>	<u>e</u>	<u>e</u>	<u>ty</u>
Acetate	(NH ₄)HCO ₃	No	4.7	1.0	32	56	7	8:1

Acetate	(NH ₄)HCO ₃	No	4.7	1.5	14	63	13	4.5:1
Acetate	(NH ₄)HCO ₃	No	4.7	2.0	0	78	13	6:1
Acetate	(NH ₄)HCO ₃	No	4.7	2.4	0	76	17	4.5:1
Acetate	(NH ₄)HCO ₃	No	4.7	3.0	0	71	23	3:1
Citrate	0.1% TFA	Yes	4.7	4.3	4.3	89.6	4.1	22:1
Carbonat	0.1% TGA	Yes	9.0	0.99	13	85	0	51:1**
e								

It was determined that many of the side reactions occurred during size exclusion chromatography in the ammonium bicarbonate solvent used to separate reaction products. Using an acidic solvent system, 0.1% TFA, virtually eliminated multiply labeled side products leading to considerable improvements in both selectivity and yield. Including a mild reducing agent, tris(2-carboxyethyl)phosphine (TCEP), in the reaction buffer also significantly curtailed several minor side reactions revealed by HPLC, especially disulfide formation. Labeling and gel filtration under these optimized conditions produced a 70% recovered yield of labeled SA-test peptide (90% yield based on HPLC quantitation), 90% of which was labeled with only a single dye at the aminooxy amine. The labeling selectivity was increased to 22/1 (Table 1).

The purity of the product was confirmed using RP-HPLC, mass spectroscopy, further chemical reaction, and isolation of purposefully overlabeled products. The labeled peptide eluted as a single peak under all HPLC conditions tested with a mass consistent with that predicted for singly-labeled SA-test peptide (Mass=1970). To determine that the labeling site was indeed at the N-methylaminooxy group, a selective zinc/acetic acid reduction procedure was used to cleave the N-O bond (Figure 3). HPLC of the reduction reaction showed >98% conversion of the starting material and a new earlier eluting peak. The mass of this peak (1530 amu) corresponded to the predicted mass of the unlabeled SA-test peptide cleaved at the aminooxy N-O bond. The residual zinc was washed several times with a saturated solution of EDTA in water, which demonstrated that the reduction reaction was complete.

To eliminate the unlikely possibility that the HPLC peak containing isolated single-labeled SA-test peptide product was a mixture of two

labeled species, SA-test peptide was reacted under higher pH conditions (pH 9.0) to label all reactive sites. SA-test peptide contained 3 nucleophilic labeling sites which would be irreversibly labeled: aminooxy, lysine, and N-terminal amine. Dye labeling at high pH generated a mixture of peptides labeled at all possible combinations of sites with 1, 2 or 3 dyes. HPLC analysis of this reaction mixture showed 8 peaks, indicated by ESI-MS to correspond to unreacted SA-test peptide, SA-test peptide single-labeled at the aminooxy nitrogen, and six additional peaks corresponding to two single-labeled peptides, three peptides bearing two dyes and a single triply-labeled peptide species. This experiment revealed the HPLC retention times of all these products, none of which co-eluted with the peak identified as SA-test peptide labeled with a single dye on the secondary aminooxy nitrogen.

Site-specific labeling of the aminooxy group could even be achieved at basic pH (Table 1). Using pH 9.0 carbonate buffer in our solvent system, addition of 0.5 equivalents of dye produced, after 3 hours, ~50% conversion of the starting SA-test peptide to a single peak with the elution time of the desired singly-labeled product. After addition of another 0.5 equivalents of TMR-OSu and an additional 3 hours of reaction time, HPLC showed ~85% conversion to a peak with the retention time of the desired product. Two minor peaks (~2-3% of total peak area), were also apparent and corresponded to the two other single-labeled SA-test peptide species identified in the multiple labeling experiment above. The N-hydroxysuccinimide ester of rhodamine clearly showed selective reactivity with the aminooxy group.

The selectivity observed at higher pH cannot be explained by the nucleophilicity of the aminooxy group alone. In fact, others have shown that at, in an uncatalyzed reaction with phenyl acetate at high pH, amines are more reactive than O-alkyl aminooxy groups. Therefore we suggest that kinetic factors are contributing to the selective reactivity of the N-methylaminooxy group, even when competing groups are not protonated. Possible reasons for this include: (1) the aminooxy oxygen localizes the nitrogen near the activated ester via formation of a hydrogen bonded "bridged" intermediate (2) a base catalyzed reaction pathway under the conditions of our reaction. This exceptional reactivity has important practical implications, as it can allow the selective

labeling of acid-labile polypeptides and synthetic proteins under physiological or basic conditions.

Example 2. The secondary aminooxy group is compatible with C^α-thioesters and amide-forming ligation.

Preparation of proteins by total chemical synthesis often requires the ligation of large polypeptides prepared by solid-phase peptide synthesis on thioester-linker resins. The most generally applicable methods available for ligations are native chemical ligation and expressed protein ligation. These processes utilize the same basic chemistry to join two peptides, one with an N-terminal cysteine and the other with a C-terminal thioester, through a regiospecific and site-specific reaction to generate a larger polypeptide. The application of aminooxy-labeling chemistry to the synthesis of large polypeptides and proteins requires compatibility with these solid-phase peptide synthesis and ligation chemistries.

The optimal approach for utilizing aminooxy-labeling chemistry in the chemical synthesis of proteins is direct incorporation of the aminooxy group as part of an amino acid used during standard solid-phase peptide synthesis. For this purpose, we generated a suitably protected N-methylaminooxy amino acid, α -Boc- β -[N-(2-Chlorobenzoyloxycarbonyl)-N-Methylaminooxy Acetyl]- α,β -Diaminopropionic Acid [Boc-2-Cl-Z-(SA)Dapa-OH] (4), as shown in Figure 4. This amino acid, referred to as SAOD, was incorporated into the peptide sequence LY-(SAOD)-AG-MPAL thioester by synthesis on TAMPAL thioester-linker resin, as described below in the Methods. (MPAL is the C-terminal mercaptopropionyl-leucine group generated by cleavage of a peptide from TAMPAL resin, see Hojo, H., et al., *Bull. Chem. Soc. Jpn.* 1993, 66:2700-2706; and Hackeng, T.M. et al., *Proc. Natl. Acad. Sci. USA*. In press)

Ligation of the LY-(SAOD)-AG-MPAL (SEQ ID NO: 3) thioester peptide with the peptide CRANK-NH₂ (SEQ ID NO: 4), was tested using standard procedures employing phosphate buffer with 6M guanidine hydrochloride at neutral pH in the presence of 2-3% thiophenol by volume. The ligation proceeded over 24 hours and generated the desired ligation product, LY-(SAOD)-AGCRANK-NH₂ (SEQ ID NO: 5), at ~85% yield. The major side

product was attributable to modification of unligated CRANK-NH₂ peptide under the ligation conditions (mass=714.5, data not shown), and was not related to the presence of the aminooxy group. There was also a single time-dependent side reaction, which generated a product of 14 mass units lower than the desired.

- 5 Using high concentrations of reacting peptides and isolating the ligation product after 24 hours reduced this side reaction to acceptable levels (<5%).

The ligation of a peptide containing multiple potentially reactive functional groups, including a hexahistidine tag useful for affinity chromatography was also tested. Coupling

- 10 CEYRIDRVRLFVDKLDNIAQVPRVGAA-HHHHHH (SEQ ID NO: 6) to LY-(SAOD)-AG-MPAL thioester proceeded to completion in 5 hours with minimal side reactions. In both ligation reactions, there was less than 1% LY-(SAOD)-AG-MPAL self condensation product, indicating that the aminooxy group and thioester do not appreciably react with one another under the ligation conditions.
- 15 These results demonstrate that the inclusion of an unprotected aminooxy group in the peptide chain is compatible with native chemical ligation.

- Labeling of the two ligation products using tetramethylrhodamine succinimide ester proceeded with selectivity similar to that for the SA-test peptide. HPLC integration indicated that the product of the LY-(SAOD)-AG-
- 20 MPAL ligation with CRANK-NH₂ was labeled with greater than 95% efficiency and with a selectivity of 34:1. Mass spectral analysis and zinc reduction demonstrated labeling at only the aminooxy group. For the longer hexahistidine-containing polypeptide ligation product, selectivity for the aminooxy group was greater than 10:1, but it was difficult to achieve high yields. The histidines could
- 25 potentially have been affecting yield and selectivity by catalyzing nucleophilic attack on the succinimide ester of the reactive dye. Inclusion of guanidine hydrochloride in the reaction solvent increased the yield to approximately 50%, indicating that folding or poor solubility of the peptide was a factor in preventing access of the reactive dye to the aminooxy group. Selectivity was also
- 30 improved, presumably because of the availability of the reactive secondary aminooxy group. Single-site labeling at the aminooxy group was proven by mass spectral analysis of trypsin and α -chymotrypsin digests of the labeled polypeptide product.

Example 3. Specificity of labeling in protein domains containing aminooxy amino acids.

As a control to establish the selectivity of labeling for aminooxy amino acid, the labeling of native β -Lactoglobulin with tetramethylrhodamine N-hydroxysuccinimide ester was attempted. It was found that non-specific labeling of this 162 aa protein containing 15 lysines and 4 cysteines was minimal (<1%), even after 6 hours.

Finally, the GTPase binding domain of p21 activated kinase (45 aa, 4 lys, 1 cys) with a secondary aminooxy amino acid incorporated at the amino-terminus (SAOD-PBD) was prepared. Previous experiments have demonstrated that PBD domains labeled with fluorescent reporter dyes at this terminus could be used as biosensors of GTPase activation. Labeling of PBD using the new methodology would enable the production of sufficient quantities to apply the biosensors *in vivo* and in pharmaceutical screening applications, and would allow incorporation of sensitive detectable groups enabling applications within living cells.

SAOD-PBD was readily labeled with Alexa-532 N-hydroxy-succinimide ester by titration addition of dye at pH 4.7 over 72 hours. The labeling efficiency was commensurate with that reported for the longest model peptide (~50% yield by HPLC quantitation) and there was no indication of multiple labeling. In this case, isolation of labeled SAOD-PBD by RP-HPLC proved difficult. Separation to baseline resolution was not achieved, but small quantities of unlabelled PBD in the labeled product do not preclude the use of the labeled material in biosensor applications. Previous reports indicate that separation of labeled product from starting polypeptide is highly dependent on the specific peptide and the attached dye.

These results demonstrate that the optimized site-specific labeling chemistry reported here is compatible with the steps required for the preparation of proteins by total chemical synthesis.

Example 4. Materials and Methods.

General: For column chromatography, silica gel (230–400 mesh) was used in standard glass columns with gravity or air pressure. Reversed-phase high performance liquid chromatography (RP-HPLC) was performed on a Waters HPLC system with UV detection at 214 nm using either a Vydac C-18 analytical column (5 μ m, 0.46 x 25cm), a Waters RCM 8 x 10 module equipped with a semi-preparative Delta Pack C-18 Radial Pack cartridge column (15 μ m, 8 x 100 mm) from Millipore, or a Vydac C-18 preparative scale column (15 μ m, 1.0 x 25 cm). Linear gradients of solvent B (0.09% TFA in 90% acetonitrile/10% water) in solvent A (0.1% TFA in water) were used for all HPLC chromatographic separations.

Mass spectra of peptides were obtained either with a Sciex API-III electrospray ionization (ESI) triple-quadrupole mass spectrometer (PE Biosystems, Foster City), or matrix assisted laser desorption ionization time-of-flight (MALDI-TOF) instruments from Thermo Bioanalysis (Thermo Bioanalysis, LTD., UK) or Kratos Analytical (Chestnut Ridge, NY). For ESI-MS, the observed masses reported were derived from the experimental m/z states for all observed charge states of a molecular species using the program MacSpec (Sciex, version 2.4.1) for electrospray mass spectrometry. MALDI-MS observed masses were relative to internal calibration using α -cyano-hydroxycinnammic acid or sinipinic acid matrices. Calculated masses reported were derived from either MacProMass (Terry Lee and Sunil Vemuri, Beckman Research Institute, Duarte, CA) or PAWS (Version 8.1.1, ProteoMetrics) and reflect the average isotope composition of the singly-charged molecular ion. Proton nuclear magnetic resonance spectrometry was recorded on a Bruker AC-250 mass spectrometer and data was analyzed using WinNMR (Bruker Instruments). Ultraviolet-Visible spectroscopy was performed on a Hewlett-Packard photodiode-array spectrophotometer.

Boc-L-amino acids were purchased from Novabiochem (La Jolla, CA) or Bachem Bioscience, Inc. (King of Prussia, PA). [[4-(Hydroxymethyl)phenyl]-acetamido]methyl (-OCH₂-Pam) Resin was purchased from PE Biosystems (Foster City, CA) and methylbenzhydrylamine (MBHA) resin was purchased from Peninsula Laboratories, Inc. (San Carlos, CA). Solvents were Synthesis

grade or better and were purchased from Fisher Scientific (Tustin, CA).
 Trifluoroacetic acid (TFA) and anhydrous hydrogen fluoride were purchased
 from Halocarbon (New Jersey) and Matheson Gas (Rancho Cucamonga, CA).
 Dyes were obtained from Molecular Probes (Eugene, OR). All other reagents
 5 were analytical grade or better and were purchased from Aldrich (Milwaukee,
 WI), Lancaster (Windham, NH), Peptides International (Louisville, KY) or
 Richelieu Biotechnologies (Montreal, Canada).

Peptide Segment Synthesis. Synthesis of peptides was carried out manually
 10 using optimized stepwise solid-phase synthesis methods with *in situ*
 neutralization and HBTU activation procedures for Boc chemistry on either -
 OCH₂-Pam, MBHA, or Trt-protected mercaptopropionyl-Leu (TAMPAL) resin
 (Hojo, H., et al., *Bull. Chem. Soc. Jpn.* 1993, 66:2700-2706; Hackeng, T.M. et
 al., *Proc. Natl. Acad. Sci. USA*. In press; and Schnolzer, M., et al., *Int. J.*
 15 *Peptide Protein Res.* 1992, 40:180-193). Standard Boc protecting group
 strategies were employed. Coupling was monitored by quantitative ninhydrin
 assay after 15 minute coupling cycles. After chain assembly, standard
 deprotection and cleavage from the resin support was carried out by treatment at
 0 °C for 1 hour with anhydrous HF containing either 10% p-cresol or anisole as
 20 scavenger. Purification was performed using RP-HPLC.

**Synthesis of TAMPAL Resin (Hojo, H., et al., *Bull. Chem. Soc. Jpn.* 1993,
 66:2700-2706).** 2.5 grams of MBHA resin (0.865 mmol/g, 2.16 mmol of amine)
 was swelled in DMF. Boc-Leu-OH (1.1 grams, 4.4 mmol) was activated with
 25 HBTU (8 ml, 0.5M solution) and DIEA (2 ml), then coupled to the MBHA
 resin until complete reaction by ninhydrin assay. The N^α-Boc group of the
 linked leucine was removed with neat TFA, then S-Trt-β-mercaptopropionic acid
 (1.5 grams, 4.3 mmol), activated in the same manner as Boc-Leu-OH, was added
 to the deprotected Leu-MBHA resin and allowed to couple until complete
 30 reaction. The S-Trt-β-mercaptopropionyl-Leu-MBHA resin was washed
 extensively with DMF, then DCM/MeOH (1/1), and finally dried *in vacuo* to
 yield 3.39 grams of thioester resin. Substitution calculated by weight gain
 yielded 0.549 mmol/gram.

Deprotection of TAMPAL Resin: S-trityl protection was removed by two 5 minute treatments with 95% TFA/5% triisopropylsilane. The deprotected resin was extensively with DMF before coupling the first amino acid, activated using optimized *in situ* neutralization protocols.

5

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylhydroxylamine (1)
(Jencks, W.P., Carriuolo, J. *J. Am. Chem. Soc.* 1960, 82:675; Jencks, W.P. *J. Am. Chem. Soc.* 1958, 80:4581, 4585). N-methylhydroxylamine

hydrochloride (0.95g, 11.37mmol) was dissolved in 3ml of water with rapid
10 stirring. The pH of this solution was adjusted to 6-7 by dropwise addition of a saturated solution of sodium bicarbonate. 2-Chlorobenzyloxycarbonyl-N-hydroxysuccinimidyl carbonate (1.2g, 4.23mmol) was dissolved in 4ml of THF and added slowly to the rapidly stirring solution of neutralized N-methylhydroxylamine. After stirring at room temperature for 14 hours, the
15 reaction was quenched with 20ml of saturated sodium bicarbonate and extracted three times with 20ml ethyl acetate. The combined ethyl acetate layers were washed once with saturated sodium bicarbonate, dried over anhydrous sodium sulfate and the solvent was removed *in vacuo* to yield 0.77g (3.77mmol, 84%) of an off-white solid. TLC R_f=0.2 (Hex/EtOAc/AcOH 80/20/1). ¹H NMR: 3.23(s,
20 3H), 5.25 (s, 2H), 7.24 (m, 2H), 7.37 (m, 2H). HRMS: Expected=216.0427, Observed=216.0425.

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylaminoxy Acetic Acid-Tert-butyl ester (2) (Jerry March *in Advanced Organic Chemistry, Third*

25 *Edition. John Wiley & Sons, New York. 1989, pp381; and Nyberg, D.D., Christensen, B.E. J. Am. Chem. Soc.* 1957, 79:1222; Motorina, I.A., et al., *Synlett* 1996, 389). Compound 1 (0.96g, 4.71mmol) was dissolved at room temperature in 10ml of THF with rapid stirring. Bromoacetate tert-butyl ester (1.05g, 5.38mmol) was added, then sodium iodide (1.5g, 10.01mmol) followed
30 by DIEA (2.5ml, 15.92mmol). The reaction changed to an orange-yellow color after addition of sodium iodide. The reaction was quenched with 30ml water after complete reaction (~3 hours) and extracted 3 times with ethyl acetate. The combined ethyl acetate layers were dried over sodium sulfate and the ethyl acetate was removed *in vacuo*. The resultant oily solid was purified by silica

chromatography on 230-400 mesh silica gel using hexanes/ethyl acetate/acetic acid (80/20/1) to yield 1.40g (4.29mmol, 90%) of a pure yellow oil. TLC Rf=0.5 (Hex/EtOAc/AcOH 80/20/1). ¹H NMR: 1.46 (s, 9H), 3.29 (s, 3H), 4.36 (s, 2H), 5.26 (s, 2H), 7.24 (m, 2H), 7.40 (m, 2H). HRMS: Expected Mass=330.1108, 5 Observed Mass=330.1104.

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylaminoxy Acetic Acid (3) (Bryan, D.B., et al., *J. Am. Chem. Soc.* 1977, 99:2353). Compound 2 (1.1g, 3.30mmol) was dissolved into 4ml of DCM and, with rapid stirring, neat TFA 10 (5ml) was added dropwise over 2 minutes at room temperature. After 1 hour, the reaction was quenched with 20ml water, extracted 3 times with DCM, and the combined DCM layers were dried over sodium sulfate. The DCM was removed *in vacuo* to yield 0.9g (3.28mmol, 99%) of an off-white solid. TLC Rf=0.2 (Hex/EtOAc/AcOH 80/20/1). ¹H NMR: 3.23 (s, 3H), 4.50 (s, 2H), 5.32 (s, 2H), 15 7.28 (m, 2H), 7.40 (m, 2H). HRMS: Expected Mass=274.0482, Observed Mass=274.0479.

Synthesis of N-(2-Cl-benzyloxycarbonyl)-N-Methylaminoxyacetyl- α -Boc- α,β -Diaminopropionic Acid [(SA)Dapa-OH](4) (Wahl, F., Mutter, M. *Tett. Lett.* 1996, 37:6861-6864; and Anderson, G.W., et al., *J. Am. Chem. Soc.* 1964, 86:1839). N-(2-Chlorobenzyloxycarbonyl)-N-Methylaminoxy Acetic Acid (3) (2.5g, 9.2mmol) was activated with N-hydroxysuccinimide (2.11g, 2equiv.) and DIC (1.440ml, 1.0equiv.) in 20ml DCM. This reaction was rapidly stirred at room temperature for 2 hours prior to the addition of N ^{α} -Boc- α,β - 25 diaminopropionic acid (2.3g, 1.2equiv.) and DIEA (3.20ml, 2equiv.). After 4 hours, the DCM solvent was removed *in vacuo*, and 50ml ethyl acetate was added. The ethyl acetate layer was washed twice with 0.5M acetate buffer, pH=4.0, then twice with 0.1N sulfuric acid. The combined acid washes were then washed with 50ml ethyl acetate. The combined ethyl acetate layers were 30 dried over sodium sulfate, then concentrated *in vacuo* to yield a viscous yellow oily solid. This solid was subjected to 3 hexane precipitations from diethyl ether to yield 2.16g (51% yield) of an off-white solid. TLC Rf=0.2-0.4 (Hex/EtOAc/AcOH 30/70/0.5). ¹H NMR: 1.45 (s, 9H), 3.16 (s, 3H), 3.54 (d-of-t, 1H, J=14.3, 4.6Hz), 3.93 (m, 1H, J=14.3, 7.5, 4.6Hz), 4.31 (s, 0.5H), 4.38 (s,

1H), 4.45 (s, 1H), 4.51 (s, 0.5H), 5.34 (s, 2H), 5.97 (broad-d, 1H, J=7.3Hz), 7.30 (m, 2H), 7.43 (m, 2H), 8.50 (broad-s, 1H). HRMS: Expected Mass=460.1487, Observed Mass=460.1480.

- 5 **Synthesis of Secondary Aminooxy Test Peptide (SA-test peptide).** The SA-test peptide, NH₂-AKAARAAAK*AARACA-CO₂H, was synthesized with Lys 10 side chain Fmoc protection as described previously (Canne, L.E., et al., *J. Am. Chem. Soc.* 1995, 117:2998-3007). Incorporation of the secondary aminooxy group was accomplished by coupling 2-Cl-Z protected N-methylaminooxyacetic (300mgs, 1.09mmol) activated with
- 10 Diisopropylcarbodiimide (157ul, 1.00mmol) and N-hydroxysuccinimide (140mgs, 1.22mmol) in 2 ml DCM for 1-2 hours, then diluted with 2 ml DMF just prior to coupling to the ε-amino group of Lys 10. Optimized coupling, cleavage and purification protocols were utilized. Amino acid analysis was
- 15 consistent with the desired peptide. Expected Mass=1560, Observed Mass=1559.

- Synthesis of LY-(SAOD)-AG-MPAL-Thioester.** LY-(SAOD)-AG-MPAL-Thioester was synthesized using optimized *in situ* neutralization protocols for
- 20 Boc chemistry on TAMPAL resin. Coupling of the N^α-Boc-(SA)Dapa-OH amino acid was accomplished by reacting the *in situ* activated N-hydroxysuccinimide ester to the deprotected amino-terminal nitrogen of alanine (Canne, L.E., et al., *J. Am. Chem. Soc.* 1995, 117:2998-3007). (SA)Dapa-OH (4) (230mgs, 0.5mmol) was dissolved in 1ml DCM and N-hydroxysuccinimide
- 25 (115.1mgs, 1.0mmol) and DIC (74.4μl, 0.47mmol) were added. The reaction was mixed briefly and allowed to activate for 1-2 hours at room temperature prior to coupling to the deprotected N-terminus of the peptide chain. After this coupling, no further modifications of the synthetic protocols were required. Expected mass=797, Observed mass=797.

30

Ligation of LY-(SAOD)-AG-MPAL-Thioester with CRANK-NH₂ Peptide. LY-(SAOD)-AG-MPAL-Thioester (3mg, 3.8μmol) was dissolved into 100 ul of 50mM phosphate buffer containing 6M guanidine hydrochloride, pH=7.2. To

this solution was added CRANK-NH₂ peptide dissolved into 100μl of the same phosphate buffer and 3ul of thiophenol. The reaction was monitored by analytical reversed-phase HPLC. After 24 hours, the ligated product, LY-(SAOD)-AGCRANK-NH₂ (SEQ ID NO: 10), was isolated by semi-preparative reversed-phase HPLC (gradient=10-50% B over 60 minutes) and lyophilized to yield a fluffy white solid. Amino acid analysis was consistent with the desired product peptide. Expected Mass=1168, Observed Mass=1168.

Ligation of LY-(SAOD)-AG-MPAL-Thioester with

- 10 **CEYRIDRVRLFVDKLDNIAQ-VPRVGAA-HHHHHH (SEQ ID NO: 7).** LY-(SAOD)-AG-MPAL (0.3mgs, 0.37mmol) and CEYRIDR-VRLPFVDKLDNIAQ-VPRVGAA-HHHHHH (1.5mgs, 3.8mmol) were subjected to the same ligation and purification conditions as described above to yield 1.0mgs (58% yield) of a white fluffy solid. Expected Mass=4518, 15 Observed Mass=4517.

Synthesis of Amino-Terminal P21 Binding Domain (PBD) Peptide

Fragment, (SAOD)-KKKEKERPEISLPSDFEHTIHVGFDA-MPAL

- Thioester (SEQ ID NO: 8):** Secondary aminooxy containing amino-terminal PBD thioester were synthesized as described above using TAMPAL resin. HF 20 cleavage utilizing p-Cresol scavenger followed by HPLC purification yielded (SAOD)-KKKEKERPEISLPSDFEHTIHVGFDA-MPAL containing two DNP groups protecting the histidines. Mass Expected=3745, Mass Observed=3745.

25 **Synthesis of Carboxy-Terminal of P21 Binding Domain (PBD) Peptide**

Fragment, CTGEFTGMPEQWARLLQT (SEQ ID NO: 9): The native carboxy-terminal half of PBD was synthesized using standard Fmoc synthesis protocols by the Scripps Peptide and Protein Core Facility. Mass Expected=2068, Mass Observed=2068.

30

Synthesis of SAOD-Modified PBD, SAOD-KKKEKERPEISLPSDFEHTIH-VGFD
ACTGEFTGMPEQWARLLQT (SEQ ID NO: 11): 1.5 mg of SAOD-KKKEKERPEISLPSDFEHTIHVGFDAMPAL (0.4 mmol) was ligated to 1 mg (4.8 mmol) of carboxy-terminal fragment, CTGEFTGMPEQWARLLQT, as

described above. After 48 hours, the ligated PBD proteins were isolated by RP-HPLC and lyophilized, 1.5mg (70% yield), Mass Expected=5262, Mass Observed=5262.

- 5 **Selective Labeling of SA-Test Peptide with Tetramethylrhodamine N-hydroxysuccinimidyl ester.** A solution of SA-test peptide (3.396µg/µl, 2.18mM) in 5% acetate buffer, pH=4.7 incorporating 5mM TCEP was utilized for labeling. A stock solution of dye (5 µg/µl, 9.5 mM) was made by dissolving TMR-OSu in neat DMSO. For each reaction, the dye stock was diluted so that
- 10 the desired number of dye equivalents could be added in 20 µl of DMSO. The following equivalents of dye were tested: 1.2, 1.5, 1.8, 2.0, 2.4, 3.0, and 4.3. With constant stirring, 20µl of dye solution was added in two 10µl aliquots to 20µl of peptide solution at room temperature. The second aliquot of dye in DMSO was added 10 minutes after the initial dye addition. After complete
- 15 addition of dye, the reaction was briefly vortexed, then incubated at room temperature. After 3 hours, the labeled reaction product(s) were separated from unreacted dye by gel filtration on Sephadex G-10 or G-15 columns using either 100µM ammonium bicarbonate or, after optimization, 0.1% TFA in water. The individual peptide product(s) were then separated by RP-HPLC and analyzed by
- 20 ESI-MS. Mass Expected=1971, Mass Observed=1970.

- Non-Selective Labeling of SA-Test Peptide with Tetramethylrhodamine N-hydroxysuccinimide ester.** SA-test peptide (3.396µg/µl, 2.18mM) was dissolved 100mM sodium carbonate, pH=9.01 containing 5 mM TCEP. 18µl of
- 25 neat DMSO was added to 20µl of peptide solution. With rapid mixing, 1µl stock dye solution in DMSO was added to this mixture. Upon completion of addition, the reaction was vortexed briefly, then incubated at room temperature. After 3 hours, a 15µl aliquot was removed and evaluated by RP-HPLC and ESI-MS. This process was repeated until significant levels of reaction were detected by
- 30 formation of single-labeled SA-lysine test peptide products.

Selective Labeling of LY-(SAOD)-AGCRANK-NH₂ and LY-(SAOD)-AGCEYRIDRVR-LFVDKLDNIAQVPRVGAA-HHHHHH with

Tetramethylrhodamine N-hydroxy-succinimidyl Ester. A sample of 20 μ L of LY-(SAOD)-AGCRANK-NH₂ (2.5 μ g/ μ L, 2.18mM) in 200mM citrate buffer, pH=4.7, 5mM TCEP was labeled using 4.3 equivalents of dye in DMSO, purified and analyzed. Expected Mass=1580, Observed Mass=1580. A 20 μ L sample of LY-(SAOD)-AGCEYRIDRVRLFVDKLDNIAQVPRVGAA-HHHHHH (SEQ ID> NO: 12) (9.7 μ g/ μ L, 2.12mM) in 200mM citrate buffer, pH=4.7, containing 5mM TCEP and 3M guanidine hydrochloride was labeled using a modified procedure. 18 μ L of a solution of tetramethylrhodamine N-hydroxysuccinimide (10 μ g/ μ L) in DMSO was added in 6 μ L aliquots over 15 minutes with rapid mixing. The reaction was incubated at room temperature for 5 hours prior to gel filtration/RP-HPLC purification and mass spectral analysis. Expected Mass=4930, Observed Mass=4929.

Labeling of β -Lactoglobulin with Tetramethylrhodamine N-

hydroxysuccinimide ester: 10 μ L of a 10 mg/ml solution of tetramethylrhodamine N-hydroxysuccinimide ester in DMSO (9.5 equivalents compared to protein), was added to a solution containing 10 μ L of DMSO and 20 μ L of a solution of β -Lactoglobulin (20.3 mg/ml or 1.1 mM) in 2.8 M guanidine hydrochloride with 5 mM TCEP (pH 4.7). After 3 hours, protein was separated from unreacted dye by gel filtration, and labeling was determined by analysis of the dye-to-protein ratio (protein concentration was determined by the method of Waddel and ϵ for tetramethylrhodamine in phosphate buffer, pH=8.0 of 81,000).

Labeling of PBD Proteins with Alexa-532 N-Hydroxysuccinimide Ester:

150 μ g of the SAOD-modified PBD protein was dissolved in 105 μ L of 200 mM sodium citrate buffer, pH=4.8, containing 5 mM TCEP (protein concentration \sim 0.28 mM). A solution of Alexa-532-OSu in DMSO (dye concentration \sim 10 mg/ml in DMSO) was titrated into the protein solution in 5 μ L aliquots over 72 hours. Four hours after each addition, the extent of labeling was determined by RP-HPLC and MS. Labeling was continued until quantities of double-labeled PBD was obtained (SAOD-modified PBD). Alexa-532 labeled SAOD-modified PBD, Mass Expected=5871, Mass Observed=5870.

Zinc/Acetic Acid Reduction of the N-methylaminoxoy N-O Bond in

Peptides. Reductive cleavage of the N-O bond was performed using zinc and aqueous acetic acid. Effervescence in the reaction was evident after a few

- 5 seconds and subsided after 60-120 minutes. After 14 hours, the reaction supernatant was analyzed by RP-HPLC and ESI-MS. Reduction of labeled SA-Test peptide, Expected Mass=1530, Observed Mass=1530: Reduction of labeled (SAOD)-AGCRANK-NH₂: Expected Mass=1139, Observed Mass=1139.

10

Trypsin/Chymotrypsin Cleavage of Labeled LY-(SAOD)-

AGCEYRIDRVRLFVDKLDNIAQVPRVGAAHHHHHHH Peptide. 10 µl of a 0.05 mg/ml solution of either trypsin or α-chymotrypsin in 25 mM ammonium carbonate (without pH adjustment) was added to 5 µl of a 10-20 µg/µl solution
15 of pure tetramethylrhodamine-labeled peptide in water (final concentration of protease is 0.033 mg/ml). The reaction was incubated at room temperature for 24 hours prior to analysis of peptide fragments by MALDI-MS.

Example 5: Synthesis and Characterization of Fluorescent Dyes

- 20 *Materials and Methods.* Commercially available solvents and reagents were purchased from major suppliers. ¹H NMR spectra of dye solutions in DMSO-d₆ were recorded at 500.11 Hz on Bruker Avance DRX-500 MHz. The pentet corresponding to the residual protons of DMSO-d₆ (2.49 ppm) was used as internal reference.

- 25 The analytical sample of the dyes was purified by HPLC on a Vydac C-18 column (no. 218TP152022, 22*250 mm, 15-20 µm, 3mL/min) using acetonitrile-water gradient elution.

Absorbance spectra were recorded on Hewlett-Packard HP8452 diode array spectrophotometer and fluorescence spectra were recorded on SPEX

- 30 Fluorolog 2 spectrometer.

All reactions were run in a oven-dried round bottom flask, under argon atmosphere, and capped with a rubber septum. 1-Benzothiophen-3(2*H*)-one 1,1-dioxide was synthesized according the procedure of M. Regitz (Chem. Ber., 1965, **98**, 36). 3-(dicyanomethylene)indan-1-one was synthesized from 1,3-

indandione by the method of K. A. Bello, L. Cheng and J. Griffiths (Chem. Soc. Perkin Trans. II, 1987, 815-818).

Preparation of (2Z)-2-[(2E)-3-METHOXYPROP-2-ENYLIDENE]-1-BENZOTHIOPHEN-3(2H)-ONE 1,1-DIOXIDE (Compound 1-SO),
(2Z)-2-[(2E)-3-METHOXYPROP-2-ENYLIDENE]-3-(DICYANOMETHYLENE)INDAN-1-ONE (Compound 1-CN).

To a solution of corresponding ketone (2.5 mmol) in 20-50 ml of methanol was added trifluoroacetic acid in one aliquot at 80°C followed immediately by the addition of 2 ml of 1,3,3-trimethoxy propene. The reaction mixture turned into a clear dark brown solution and after an interval of 30 sec a yellow solid precipitated out. The precipitate was filtered and dried.

For compound 1-SO, the yield was 60%. ¹H NMR (CDCl₃): δ 4.04 (s, 3H, CH₃O), 6.53 (t, ³J_{H-H} = 12.1Hz, 1H), 7.53 (d, ³J_{H-H} = 11.7 Hz, 1H), 7.74 (d, ³J_{H-H} = 12.8 Hz, 1H), 8.2-8.4 (m, 4 H, Ph).

For compound 1-CN, the yield was 60%. ¹H NMR(CDCl₃): δ 8.67 (dd, 6.2 Hz, 2.9 Hz, 1H), 8.37 (d, 11.7 Hz, 1H), 7.87 (m, 1H), 7.73 (m, 3H), 7.51 (d, 12.05 Hz, 1H), 3.98 (s, 3H).

PREPARATION OF (2Z)-2-[(2E,4E)-4-(3-(3-SULFONATOPROPYL)-1,3-BENZOTHAZOL-2(3H)-YLIDENE)BUT-2-ENYLIDENE]-1-BENZOTHIOPHEN-3(2H)-ONE 1,1-DIOXIDE (DYE 2).

1.16 g (2.00 mmol) of 3-(2-methyl-1,3-benzothiazol-3-ium-3-yl)propane-1-sulfonate and 625 mg (2.50 mmol) of compound 1-SO were dissolved in 65 ml of mixture CH₃OH-CHCl₃ (1:1) and heated at reflux. 328 mg (4.00 mmol) of AcONa in 10 ml was added at once and the reaction mixture was refluxed for additional 30 min. After cooling, the dye was separated by filtration, recrystallized from methanol and dried. The yield was 65%.

¹H NMR: (DMSO-d₆) δ 1.85 (p, ³J_{H-H} = 6.2 Hz , 2H, CH₂-CH₂-CH₂), 2.60 (t, ³J_{H-H} = 6.2 Hz , 2H, CH₂-CH₂-SO₃), 4.45(t, ³J_{H-H} = 6.2 Hz , 1H, CH₂-N), 6.87 (d, ³J_{H-H} = 13.2 Hz, 1H), 7.3 - 8.1 (m, 11H).

Table 2. Selected Merocyanine Dyes.

Compound	Dye Structure
1	
2	
3	
4	

Table 3. Spectral Data for Selected Merocyanine Dyes.¹

Compound	Extinction coefficient (*10 ⁻³)	Absorption Maximum	Emission Maximum	Quantum Yield
1	100	618	636	0.07
2	100	603	623	0.10
3	120	586	618	0.17
4	80	667	693	0.10

¹All spectral data are recorded in n-butanol.

**Example 6: Imaging the Spatio-temporal Dynamics of
Rac Activation *in vivo* with FLAIR**

FLAIR (Fluorescent Activation Indicator for Rho GTPases) is a biosensor system that maps the spatial and temporal dynamics of Rac activation in living cells. The approach is based on microinjection of a fluorescently labeled domain from p21-activated kinase into cells expressing GFP-Rac. The injected domain (called PBD, for p21-binding domain) binds only to Rac-GTP, and not to Rac-GDP (Thompson et al., 37 Biochemistry 7885 (1998); Benard et al., 274 J. Biol. Chem. 13198 (1999)). Within living cells, PBD binds to the GTP-Rac wherever it has bound GTP, bringing the Alexa546 dye on the PBD near the GFP on the Rac to produce fluorescence resonance energy transfer (FRET). Thus, the FRET signal marks subcellular locations where Rac is activated. This can be quantified to follow the changing levels and locations of Rac activation or to trace the kinetics of total Rac activation on an individual cell basis.

The labeling of PBD with Alexa, and mammalian expression vectors for expression of Rac-GFP is by any procedure, for example, as described in this application. This example provides protocols for production of pure PBD, protocols for generating cell images suitable for quantitative analysis of rac activation, and procedures and caveats for generating two types of data: images showing the spatial distribution of Rac activation within cells, and curves showing the kinetics of Rac activation in single cells.

PBD Expression and Purification. PBD is expressed in the form of C-terminal 6His fusion from the prokaryotic expression vector pET23 (Novagen). It was determined experimentally that the highest levels of expression are observed when a vector containing plain T7 promoter (not T7lac) is used in combination with a BL21(DE3) strain (not the more stringent BL21(DE3)pLysS) of *E.coli*. This system allows for "leaky" protein expression (Novagen). While the 6His tag can be cleaved from the purified protein with thrombin, it is not necessary, as the tag does not have any significant effect on probe functionality.

Competent BL21(DE3) cells (*Stratagene*) are transformed with pET23-PBD using standard procedures. See, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press 1989) After transformation, cells were plated on an LB plate containing carbenicillin. Cells do not degrade carbenicillin as quickly as ampicillin, so a higher percentage of cells retain the vector at the culture density appropriate for induction (Novagen). Five ml of LB media with 100 µg/ml carbenicillin were inoculated with a single colony of cells, and grown in the shaker at 37°C for 6-8 hours (until dense). Two ml of this are then used to inoculate 50 ml of LBcarb. The rest of the culture is diluted 1:1 with glycerol and frozen for long term storage at -80°C. The 50 ml culture is incubated in the shaker overnight at 37°C. Next morning 1-2 L of LBcarb are inoculated with the overnight culture (15-20 mL culture/500 mL media), and grown in the shaker (37°C) to OD₆₀₀ = 0.8-0.9 (about 2-3 hours). After that the cultures are briefly chilled on ice to 30-32°C, then put back in the shaking incubator turned down to 30-32°C. The protein is expressed at a lowered temperature to increase the portion of the correctly folded, soluble PBD. IPTG is added to a final concentration of 0.4-0.5 mM, and the cultures are allowed to grow for another 4-5 hours at 30-32°C (shaker). The cells are collected by centrifugation (8,000 G, 4 min), and stored as a pellet at -20 °C until use. Approximately 2.5-3 g of cells is usually obtained from each liter of culture.

Purification of PBD-6His is performed essentially as described in the Talon affinity resin manual (*Clontech*). The cells (3-5 g) are thawed in 20-30 ml of the Lysis buffer (30 mM Tris HCl, pH 7.8, 250 mM NaCl, 10% glycerol, 5 mM MgCl₂, 2 mM β-ME, 1 mM PMSF), homogenized with a spatula and sonicated (4 pulses, 10-15 sec each). T4 lysozyme and DNase are added in catalytic amounts (approximately 100 micrograms/ml lysozyme and 500 U

DNAse) to help the lysis, and the suspension is incubated on ice with periodic mixing for 30 min. The cells are then centrifuged at 12,500 rpm for 30 min, and the supernatant containing PBD is carefully transferred into a 50 mL Falcon tube.

- 5 Talon resin (1.5-2 ml) (*Clontech*) is washed twice with 10 volumes of the lysis buffer in a 15 ml Falcon tube, centrifuging in the swinging bucket centrifuge at low speed in between to separate the resin. The cell lysate is added to the 1.5 ml of washed Talon resin in a 50 mL falcon tube, and inverted or agitated gently (i.e. with an orbit shaker) for 20-30 min at r.t. The resin is then
- 10 separated by centrifugation in a swinging bucket centrifuge. The supernatant containing the unbound fraction is removed and saved for SDS-PAGE analysis. The resin is then transferred into a new 15 mL Falcon tube and washed twice (10-15 min each, r.t., orbit shaker) with 12 mL of the lysis buffer, without PMSF and β ME. The third wash is performed with lysis buffer + 10 mM imidazole
- 15 (add 1 M stock in water, kept at -20°C). After the final separation, the resin is resuspended in 2-3 mL of lysis buffer with 10 mM imidazole, and pipetted into a column (0.5 cm in diameter). The resin is allowed to sediment by gravity flow until the fluid above the resin bed is almost gone, and then another 3-5 mL of Lysis buffer with 10 mM imidazole is added to wash the column. The elution is
- 20 performed using Lysis buffer with 60 mM imidazole, and ca. 500 μL fractions are collected. PBD usually elutes in fractions 5-13 (total volume about 3-4 mL). An aliquot of each fraction is run on a 12% SDS-PAGE and the fractions containing the pure PBD are combined and dialyzed against 1 L of 25 mM NaP buffer (pH 7.3). A dialysis bag (*SpectraPor 7*), or dialysis cassette (*Pierce*) with
- 25 a molecular weight cutoff value of 3,500 kDa can be used.

- After 2-3 hours of dialysis, the bag is wiped with a KimWipe and buried in Aquacide powder (*Caibiochem*) for 15-45 min at 4°C , depending on the volume of the sample in the bag. This concentration process should be monitored carefully as complete drying may occur if the bag is left in the
- 30 Aquacide for too long. The powder is scraped gently from the bag every 10-15 min to facilitate water absorption. When the sample reaches 0.5-1.5 mL in volume (3-10-fold concentration), the Aquacide is cleaned from the bag, and the sample is carefully removed. The sample is briefly centrifuged (14,000 rpm, 2 min) to separate it from the precipitated material, and transferred into a new

dialysis bag or cassette. After the second dialysis step, the concentration of PBD is measured by taking a small aliquot (5-10 μ L) and diluting into 50 mM TrisHCl (pH 7.5-8.0) or other appropriate buffer. The extinction coefficient of PBD at 280 nm is 8,250 (estimated from the primary sequence). On average, 1.5-
 5 2 mg of PBD is obtained per liter of cell culture.

Other methods of concentrating PBD were found to be less effective. For instance, centrifugal concentrators require prolonged centrifugations, and result in nonspecific adsorption of the small PBD protein to the membrane. It is essential to perform dialysis after concentration with Aquacide. This prevents the
 10 ionic strength of the resultant protein prep from becoming too high before labeling. Low ionic strength conditions are preferable to avoid excessive precipitation of the protein during attachment of the hydrophobic dye.

Loading GFP-Rac and Alexa-PBD in cells. Cells were first transfected with
 15 GFP-Rac through nuclear microinjection. The EGFP variant (*Clontech*) was used that produced significantly brighter cells than wild type GFP (Heim et al., 6 Current Biology 178 (1996)). For microinjection of DNA and of PBD-Alexa, glass pipettes with 1.0 mm outer diameter and 0.50 mm inner diameter (*Sutter*) were pulled using a micropipette puller (*Sutter* Model P-87) to make
 20 microinjection needles with tips of approximately 0.5 μ m diameter. Rac-GEP c-DNA is injected into Swiss 3T3 fibroblasts at 200 ng/ μ L, using a constant needle pressure of approximately 100 hPa. DNA can be centrifuged prior to injection (20,000 G for 15 minutes) to prevent clogging the needle.

Cells expressing GFP-Rac were microinjected with Alexa-PBD using a
 25 microscope with optics and illumination capable of revealing the GFP fluorescence (detection sensitivity is typically improved by using higher NA objectives and brighter light sources, such as a 100W Hg arc lamp). Thus, only GFP-expressing cells need to be injected. To reduce background fluorescence during injection and the following experiment, cells are placed in 1 mL of pre-
 30 warmed Dulbecco's Phosphate Buffer Solution (DPBS) containing 1000 mg/L D-glucose, 36 mg/L sodium pyruvate and supplemented with 0.2% BSA, 1% L-glutamine and 1 % Penicillin-Streptomycin. During injection and the following experiment, cells are mounted in a Dvorak live cell chamber (*Nicholson*) preheated to 37°C, and maintained at 37°C by a heated stage (*20/20 Technology*).

The microscope can be equipped with a motorized stage and shutter controls (*Ludi*) to monitor multiple stage positions in one experiment.

Cells that are barely expressing or expressing too much GFP-Rac were ignored. The former produce FRET too weak for recording, and in the latter overexpressed Rac affects the biology of the cell. In general, we observed that cells expressing less than 300 intensity units (IU) do not display Rac-induced ruffling and altered morphology. The precise value of this cutoff will depend on the sensitivity of the imaging system, and should be determined by each lab for a relevant biological behavior.

10 A 100 μ M solution of Alexa-PBD was centrifuged at 20,000 G for 1 hour prior to injection, and then injected into the cytoplasm of cells expressing the GEP-Rac. Lowering the needle into the region just adjacent to the nucleus seems to produce the best combination of efficient injection and cell health. After the injection, cells are placed back into the 37°C incubator for 5-10 minutes
15 to recover. Alexa-PBD could potentially act as an inhibitor of Rac activity, so controls were carried out showing that, for our imaging system, up to 1000 IU of Alexa-PBD do not inhibit induction of ruffling.

Imaging Rac activation. Imaging experiments were performed using a
20 *Photometrics* KAF1400 cooled CCD camera, and *Inovision* ISEE software for image processing and microscope automation. Although filters are undergoing further optimization, the best success to date has been achieved with the following filters, designed with sharp cutoffs specifically for this purpose by the Chroma corporation: GFP: HQ480/40, HQ 535/50, Q505LP FRET: D480/30,
25 HQ610/75, 505DCLP Alexa: HQ545/30, HQ610/75, Q565LP.

The exact camera settings depend on the type of experiment being performed. When the total Rac activity within the cell is being determined, images are not generated, so spatial resolution can be sacrificed for increased sensitivity. Images are taken using 3x3 binning with exposure times of 0.1 sec,
30 0.1 sec, and 0.5 sec for GFP, Alexa and FRET respectively. When images are required, i.e. to examine the changing spatial distribution of Rac activation, 1 x 1 binning is used with exposure times of 1 sec, 1 sec and up to 5 sec for GFP, Alexa and FRET respectively. These settings depend on the sensitivity of the imaging system used, and the desired trade off between sensitivity and spatial or

temporal resolution. Settings should always be chosen not exceed the dynamic range of the camera (Berland et al., in Sluder et al. (eds.), Video Microscopy at 33 (Academic Press 1998). Motion artifacts should also be considered during imaging experiments. For fast moving phenomena, features of the cell may

5 move appreciably during the time between acquisition of the FRET and GFP images. This results in artifacts when the image is corrected for bleedthrough, as described in more detail below. Such artifacts can be prevented by reducing the time between exposures, or by using two cameras simultaneously.

When total Rac activation is being determined, a picture of GFP-Rac and

10 FRET is taken at each time point. Only one image of Alexa-PBD, usually at the initial time point, will also be required (for bleedthrough corrections as described below). In contrast, when generating images (i.e. to determine the distribution of rac activation) an Alexa-PBD image is taken at each time point. The reasons for this are discussed in the 'Image Processing' section below. If the cells are to be

15 treated with some type of stimulus, it is helpful to take a series of images prior to stimulation as controls for noise, bleaching, and other artifacts.

Image Processing. Image analysis is performed to follow the kinetics of total Rac activity within individual cells, displayed as curves of activation over time,

20 or to generate images that show the subcellular location of rac activation. The proper application of corrections is needed essential for quantitative imaging. Common image processing operations can be used for this purpose (Berland et al., in Sluder et al., Video Microscopy at 33 (Academic Press 1998). Procedures for such image processing operations will depend on the software package used.

25 The correction factors should be rigorously applied when using the FLAIR system, as FRET signals will be low relative to other sources of fluorescence in the sample. The FRET signal must purposefully be kept low, as minimum quantities of fluorescent molecules should be used to prevent perturbation of cell behavior. Hence, FLAIR procedures must be more carefully controlled than

30 procedures generally used with fluorescent probes.

The following protocols assume use of a cooled CCD camera, which typically show low levels of noise, linear response to light intensity, and little variation in response from pixel to pixel. It is valuable to use cameras and software with the greatest possible bit depth (allotment of computer memory to

each pixel to maximize the number of possible intensity gradations). This is especially important for ratio operations, which are typically performed using 12 bit images or greater. Operations are described in the order in which they should be performed:

5 **1. Registration.** For corrections applied in later steps, it is important that each of the images taken using different excitation and emission wavelengths be registered so that the cells lie atop one another, with cell edges and internal features exactly coinciding. Different image processing software will accomplish this in different ways, but manual translocation is usually
10 involved so that one image lines up with a second, fixed image. This is best accomplished by zooming in on cells and adjusting brightness and contrast to clearly see cell edges and internal features. Since the GFP-Rac signal is generally the strongest, it was used as the reference image in these experiments. When the bleedthrough corrections described below are performed, errors in registration
15 often become apparent as ‘shadow effects.’

2. Background subtraction. There are two methods commonly used for background subtraction. If the only intention is to follow the changing spatial distribution of the FRET signal over time, and if the background (in the absence of cells) remains uniform across the field of view throughout the
20 experiment, then it is sufficient to determine the background intensity in several regions of the image outside the cell. The average value of these intensities is then subtracted from each pixel in the image. This method can also be sufficient for following qualitative changes in the subcellular location of activation, but it must be used with caution. Subtle variations in background intensity across the
25 cell could be large relative to the changes observed in FRET, producing artifacts. When quantifying the kinetics of total cell Rac activity over time (see following sections), it is better to take an image of a region of the coverslip containing no cells or fluorescent debris, under the same conditions used for taking the real image. This background image is then subtracted from the real image prior to
30 further analysis. A separate background image must be taken for each type of image (GFP, FRET or Alexa) and at each time point when successive images are obtained

3. Masking. It is advisable to mask out regions surrounding the cells prior to further analysis. The edges of the cell are outlined, in most software

either manually or by eliminating all sections of the images below a certain intensity value (interactive thresholding). The regions outside the cell are thus identified and eliminated from further calculations. The precise approach for accomplishing this will depend upon the software employed. However, the mask is usually a binary image with all values within the cell = 1, and all outside = 0, and the real image is multiplied by the mask. The mask is best generated using the GEP image, which has the strongest signal and therefore the most clearly defined edges. When determining the total intensity within the cell, analysis should be carried out on the same pixels within the FRET, GFP, and Alexa images. Therefore, the same mask is applied to each image after registration, assuring that exactly the same pixels are analyzed.

4. Bleedthrough correction. During FRET it is necessary to excite the donor fluorophore while monitoring emission from the acceptor fluorophore. It is extremely difficult to design FRET filters that see only FRET emission and block all GFP emission, or block all light from Alexa excited directly rather than by FRET. To correct for “bleedthrough” of such light into the FRET image, the fluorescence filters must be characterized by taking images of cells containing GFP-Rac or AlexaPBD alone. For example, in bleedthrough correction for GFP, cells are imaged using both the GFP and FRET filter set. When observing the GFP fluorescence through FRET filters, a fixed percentage of the GFP emission will be seen. The total fluorescence intensity is determined for both the GFP and FRET images from cells containing only GFP-Rac. A GFP bleedthrough factor is computed for each cell by dividing the intensity through the FRET filters by that through the GFP filters (the ‘bleedthrough factor’ for GFP: FRET intensity/GFP intensity). This value is plotted against cell intensity for numerous cells, and a line is fit to this data to produce an accurate value of the bleedthrough factor. It is important to use background-subtracted images. The process is repeated for Alexa-PBD. When the actual experiment is performed, an Alexa-PBD, GFP-Rac, and FRET image are obtained. After background subtraction of all three images, the Alexa-PBD and GFP-Rac images are multiplied by the appropriate bleedthrough factor and subtracted from the FRET image. This is an extremely important step that must be applied carefully to prevent artifacts that appear to be regions of high FRET, especially as the magnitude of the FRET signal approaches that of the bleedthrough. It is

important not to use GFP-Rac or Alexa-PBD images that exceed the dynamic range of the camera ('overexposure') as they will not fully eliminate bleedthrough. Motion artifacts can also produce errors derived from bleedthrough corrections.

5

Production of total activation curves (bleaching correction). To determine how overall Rac activity within a cell changes over time, the total fluorescence intensity within a GFP-Rac and FRET image is determined at each time point. The intensity of the FRET image is divided by that of the GFP-Rac image. This ratio better reflects total Rac activity than does FRET intensity alone. Division by GFP-Rac normalizes out errors due to bleaching of GFP, effects of uneven illumination, and other factors affecting both the GFP and FRET signals. Because all FRET occurs through irradiation of GFP, bleaching of GFP will decrease both GFP and FRET emissions to the same extent. Therefore, the FRET/GFP ratio will be a measure of Rac activity that is not affected by bleaching. It is critical that each image be properly background subtracted and corrected for bleedthrough. Bleedthrough corrections are somewhat simplified when generating these curves, as they need not be carried out on actual images, but simply on the total intensity values derived from the images. The total intensity of the GFP-Rac image is multiplied by the GFP bleedthrough factor, and the resulting value is subtracted from the FRET intensity. An analogous operation is performed for Alexa-PBD bleedthrough. One need only obtain a single Alexa-PBD image (usually at the beginning of the time series) and use this for bleedthrough correction of all images in the time series. If Alexa-PBD is not irradiated during the experiment, its bleaching will be negligible and this single image will reflect the actual Alexa-PBD level throughout the entire time series.

Any ratio calculations are best performed using floating point operations. Large errors can be generated by software that truncates noninteger values into integers to display data as images. Such problems can be overcome by multiplying the images by a large scalar value prior to division. This value should be as high as possible without causing any pixel to exceed the bit depth of the image file. For example, a 12-bit image file can only hold values up to 4095 (2^{12}), so the constant selected must not cause the highest intensity value in the

image to exceed 4095. When operations are performed on whole cell values rather than on images, as with production of the Rac activation curves described here, it is convenient to determine total cell intensities and then perform any division operations in a floating point spreadsheet program, such as Microsoft Excel.

Examining changes in the localization of Rac activation. The sections above describe how to generate images showing the distribution of FRET within cells. A sequence of such images can be compared to show how localizations vary with time. While simple examination can suffice to show distribution changes, the overall intensity of FRET, and hence the perceived activation level, could become successively lower over time due to bleaching of the GFP. We have found that bleaching is not a serious issue for at least 20 images under the exposure conditions described here. The total GFP intensity at the beginning and end of the experiment should be examined to gauge bleaching effects. One can correct for bleaching by dividing each pixel in the image by the total intensity of GFP determined at the same time point.

Many aspects of cytoskeletal control and signaling crosstalk depend upon the localization of Rho family GTPase activation, and may depend as well on the level and duration of activation. Signaling control by the precise dynamics of GTPase activation has been suggested by indirect experiments, but has been very difficult to quantify or study using previous methods. The FLAIR system reported here can reveal Rac activation dynamics *in vivo* and can accurately report the changing activation levels within a cell.

Example 7. Cellular Localization of Rac GTPase Activation Using p21 Activated Kinase (PAK) as a Protein Biosensor

Prevailing evidence suggests that signaling proteins must be tightly regulated both spatially and temporally in order to generate specific and localized downstream effects. For Rac1 and other small GTPases, binding to GTP is a critical regulatory event that leads to interaction with downstream targets and regulates subcellular localization. Here we use FLAIR (fluorescence activation indicator for Rho proteins), to quantify the spatio-temporal dynamics

of Rac1 GTP interaction in living cells. FLAIR reveals precise spatial control of growth factor-induced Rac activation in membrane ruffles, and in a gradient of activation at the leading edge of motile cells.

- Rac is a member of the Ras superfamily of small GTPase proteins (A. Hall, Science 279, 509-14 (1998). It plays a critical role in diverse signaling pathways, including control of cell morphology, actin dynamics, transcriptional activation, apoptosis signaling, and other more specialized functions (L. Kjoller, A. Hall, Exp. Cell Res. 253 166-79 (1999). The broad range of events controlled by this GTPase requires subtle regulation of interactions with multiple downstream targets. Accumulating evidence suggests that the effects of Rac are in part controlled by regulating the subcellular localization of its activation through GTP binding. For example, Rac is known to induce localized actin rearrangements to generate polarized morphological changes (C.D. Nobes, A. Hall, J. Cell Biol. 144(6), 1235-1244 (1999). GTP exchange factors (GEFs) which regulate nucleotide exchange on Rho GTPases contain a variety of localization domains and may modulate downstream signaling from Rac (Zhou et al, J. Biol. Chem. 273(27), 16782-16786 (1998).

- Although control of Rac1 signaling through localized activation is clearly important, it is difficult to study in an intact cell. Here we develop and apply a new method based on fluorescence resonance energy transfer (FRET) which can quantify the timing and location of Rac activation through GTP binding. This method is applied to examine the hypothesis that specific and localized Rac1 activation occurs during the course of extracellular signal-induced cytoskeletal dynamics.

- The design of the Rac nucleotide state biosensor is shown schematically in Figure 5. A fluorescently labeled protein biosensor is introduced into the cell together with a biologically active GFP-fused Rac (C. Subauste et al, J. Biol. Chem. 275(13), 9725-9733 (2000). This protein biosensor is labeled with an acceptor dye capable of undergoing FRET with GFP. Since the biosensor is derived from a specific GTP-Rac target protein, it binds to GFP-Rac 1 only when the Rac1 is in its activated, GTP-bound form, and produces a localized FRET signal revealing the level and location of Rac activation. When cells expressing GFP-Rac were injected with the biosensor, we were able to simultaneously map the changing location of GFP-Rac1 and the subpopulation of GFP-Rac1

molecules in the activated, GTP-bound state. FRET is proportional to the amount of GTP binding, enabling quantitation of changing activation levels. It is also very specific, as only biosensor binding to the GFP-tagged protein will generate FRET. This approach has the potential to examine many protein states, including posttranslational modifications, conformation, and ligand binding.

The biosensor was made by fluorescently labeling a domain from p21 activated kinase I (PAK1) known to bind selectively to GTP-Rac. PBD (fragment of human PAK1, residues 65-150, with a single cysteine added in the penultimate N-terminal position) was expressed as a C-terminal 6His fusion protein from the pET23 vector (Novagen) and purified from *E. coli* strain BL21DE3 using Talon metal affinity resin (Clontech) as instructed by the manufacturer. All GFP constructs were prepared using the EGFP mutant (T. Joneson, *Mol. Cell. Biol.* 19(9) 5892-901(1999); T.T. Yang et al., *Gene* 173, 19-23 (1996)), kindly provided by Dr. Roger Tsien of the University of California, San Diego. GFP-Rac 1 fusion and wild type Rac1 for *in vitro* studies were also expressed as 6His constructs and purified using a similar procedure. Purified protein was dialyzed against 50 mM sodium phosphate (pH 7.8), and labeled with 7 equivalents Alexa 546 maleimide (Molecular Probes) at 25 degrees for 2 hours. The conjugate was purified from unincorporated dye by G25 size exclusion chromatography followed by dialysis. The dye:protein ratio was between 0.8 and 1.3, as determined from absorbance of the conjugate at 558 nm (Alexa 546 extinction coefficient 104,000 M⁻¹cm⁻¹) and 280nm (PBD, extinction coefficient 8,250 M⁻¹cm⁻¹ plus Alexa absorbance, determined as 12 % of the absorbance at 546). Protein concentration was also independently determined using a Coomassie Plus protein assay (Pierce) and SDS-PAGE calibration with known concentrations of bovine serum albumin.

The p21 binding domain (PBD, aa 65-150) has been used successfully to precipitate GTP-Rac 1 from cell lysates (V. Benard, B.P. Bohl, G.M. Bokoch, *J. Biol. Chem.* 274, 13198-204(1999). In order to produce efficient FRET, the optimum site for attachment of an acceptor dye was determined by analyzing FRET between purified GFP-Rac1 and PBD labeled with various dyes in different positions. PBD contained no native cysteines, so the site of labeling could be controlled through introduction of a single cysteine, followed by labeling with cysteine-selective iodoacetamide dyes. After considerable

optimization, the best candidate was found to be a PBD with cysteine appended to the N-terminus, labeled with commercially available Alexa 546 dye. The distance between the Alexa dye at the N-terminus of PBD and the fluorophore in GFP was calculated to be 52 Å based on the efficiency of FRET and assuming random rotation of the fluorophores ($R_0 = 51$, $n = 1/4$, $k^2 = 2/3$) (T. Nomanbhoy et al., *Biochemistry* 35, 4620-4628 (1996)). We coined the name FLAIR (fluorescent activation indicator for Rho proteins) for this new live cell imaging technique.

FRET between Alexa-PBD and GFP-Rac1 was efficient and dependent on GTP-Rac1 binding. Using fluorescence excitation wavelengths that selectively excite GFP (480 nm), fluorescence emission was monitored while maintaining a fixed concentration of GFP-Rac1 and varying AlexaPBD concentrations (see Figure 7a). Purified GFP-Rac1 (200 nM) was bound to varying concentrations of GTPγS or GDP at low magnesium by 30 min incubation at 30°C as described previously (U.G. Knaus et al., *J. Biol. Chem.* 267, 23575-23582 (1992)), using the following nucleotide equilibration buffer: 50 mM Tris HCl (pH 7.6), 50 mM NaCl, 5 mM MgCl₂, 10 mM EDTA, and 1 mM DTT. Equal volumes of Alexa-PBD in the same buffer were added to the GFP-Rac1 solution and fluorescence emission spectra (500-600 nm) were acquired at room temperature and 480 nm excitation. Spectra were corrected for dilution upon Alexa-PBD addition. Alexa-PBD concentrations were either varied as shown, or maintained at 1 micromolar when saturating Alexa-PBD was required. The spectra shown were corrected for direct excitation of the Alexa fluorophore by acquiring spectra of Alexa-PBD alone at equivalent concentrations, and subtracting these from spectra shown in Figure 7. K_d were determined by fitting to the equation: $Y = A \cdot X / (K_d + X)$. In the Figure 7, panel A inset, only points actually used in the curve fitting are shown. Higher, saturating concentrations of AlexaPBD were not used because errors from subtraction of direct Alexa excitation became larger. Binding of Alexa-PBD to GFP-Rac resulted in a change in fluorescence intensity of both donor (GFP) and acceptor (Alexa) emission. When the GFP and Alexa fluorophores were brought into close proximity, FRET caused the Alexa (acceptor) emission to increase while the GFP (donor) emission decreased. No change in emission was observed when using either unlabeled PBD or Rac1 not fused to GFP, indicating that

fluorescence changes were in fact due to energy transfer (data not shown).

Because FRET alters donor and acceptor intensities in opposite directions, the ratio of emission at these two wavelengths is a sensitive measure of the PBD-Rac1 interaction. The corrected Alexa/GFP emission ratio underwent a 4-fold change upon saturation of Rac 1 with GTP (Figure 7b). The change in emission ratio versus PBD concentration was fit to the Michaelis equation to derive an apparent K_d for PBD-Rac1 binding of $1.1 \pm 0.3 \mu\text{M}$ (Figure 7a inset). This is slightly higher than the values that have been determined for various unlabeled PAK1 fragments (E. Manser, T. Leung, H. Salihuddin, Z.-S. Zhao, L. Lim, Nature 67, 40-46 (1994); D.A. Leonard et al, Biochemistry 36, 1173-1180 (1997); G. Thompson, D. Owen, P. A. Chalk, P. N. Lowe, Biochemistry 37, 7885-7891 (1998). This may indicate that the presence of the Alexa may somewhat weaken the binding interaction. Our studies in live cells, described below, demonstrated that this affinity was sufficient for detection of Rac activation in vivo, and reversible binding is in fact desirable as it permits the biosensor to rapidly respond to changes in the location or level of Rac activation. Importantly, Alexa-PBD was shown to minimally perturb Rac-GTP binding interactions. The apparent GTP γ S dissociation constant was determined at saturating Alexa-PBD by fitting the experimental data to the Michaelis equation (Figure 7b). The derived value of $47 \pm 9 \text{ nM}$ is consistent with the previously reported value of 50 nM (L. Menard, E. Tomhave, P.J. Casey, R.J. Uhing. R. Snyderman, J.R. Didsbury, Eur. J. Biochem. 206, 537-546 (1992). This validated application of FLAIR as an indicator of biologically relevant Rac-nucleotide binding.

Much is known about the localized dynamics of actin, but it has been difficult to explore how signals can generate precisely localized actin behavior. Rac has been shown to participate in signaling cascades leading to localized polymerization in several systems, but it is unclear whether Rac activity is constrained to specific subcellular regions, or whether global activation of Rac leads to more localized activation of downstream molecules. The precise spatial correlation of localized signaling and downstream actin behavior remains completely unknown. Here, we used FLAIR to study the localization and activation of Rac in two cell systems where actin behaviors are clearly constrained to specific subcellular regions, the induction of ruffling by growth

factor stimulation of quiescent cells. and polarized cell motility induced by wounding.

We examined stimulation of quiescent Swiss 3T3 fibroblasts by serum or platelet derived growth factor (PDGF) known to initiate membrane ruffling and transcription through activation of Rac 1 (Ridley A.J., Paterson H.F., Johnston C.L., Diekmann D., Hall A., Cell 70,401-10 (1992); P.T. Hawkins, Curr. Biol. 5(4), 393-403 (1995)). For serum stimulation experiments, Swiss 3T3 fibroblasts (ATCC, passage 15-27) were plated on glass cover slips and then maintained in Dulbeccos modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS), 1% L-glutamine and 1% penicillin-streptomycin for at least 24 hours. Media was then replaced with media containing only 0.5% FCS, and cells were maintained for 42 hours. Cells were transfected by microinjecting 200 micrograms/ml GFP-Rac1 c-DNA into cell nuclei 2-8 hours prior to the experiment. The EGFP mutant was used in all experiments, cloned and expressed as previously described (C. Subauste et al, J. Biol. Chem. 275(13), 9725-9733 (2000)). Cells expressing the GFP-Rac were then microinjected with 100 micromolar Alexa-PBD, mounted in a heated chamber on a Zeiss axiovert 100TV microscope and maintained in Dulbecco's phosphate buffered saline (DPBS) (GIBCO) to reduce background fluorescence. Cells were then stimulated by replacing the media with DPBS containing 10% FCS or 50 ng/mL PDGF. Images were obtained every 30 seconds using a Photometrics PXL cooled CCD camera with 1x1 or 3x3 binning, and a Zeiss 40x 1.3 NA oil immersion objective. Fluorescence filters from Chroma were as follows: GFP: HQ480/40, HQ535/50, Q505LP; FRET: D480/30, HQ610/75, 505LP; Alexa:HQ 545/30, HQ 610/75, Q565LP. Cells were illuminated using a 100W Hg arc lamp. Exposure times for 3x3 binning were: GFP- 0.1 seconds, Alexa-PBD - 0.1 seconds, FRET -0.5 seconds. For 1x1 binning, GFP - 1 second, Alexa-PBD - 1 second, FRET - 5 seconds.

Levels of exogenous proteins had to be limited to concentrations that would not perturb cell behavior. We determined the intracellular levels of Alexa-PBD and GEP-Rac 1 that altered normal serum-induced ruffle formation (Figure 8), and kept exogenous protein below these levels throughout our studies.

Image triplets of GFP, FRET, and Alexa fluorescence were taken at each successive time point before and after stimulation. Thus, as shown in Figure 9, both the changing localizations of GFP-Rac, and the level and location of Rac activation could be monitored. Images were first background subtracted and carefully registered to ensure accurate pixel alignment. The GFP-Rac image was then thresholded, changing the intensities of all pixels outside of the cell to zero. Thresholding was based on the GFP image since it had the largest signal to noise ratio, providing the clearest distinction between the cell and background. The thresholded GFP-Rac image was used to generate a binary image with all values within the cell = 1 and all outside = 0. The FRET and Alexa-PBD images were multiplied by the binary image, assuring that exactly the same pixels were analyzed in all three images. Emission appearing in the FRET image from direct excitation of Alexa and GFP was removed by subtracting a fraction of the GFP-Rac and Alexa-PBD images from the FRET image. This fraction depended on the filter set and exposure conditions used. It was determined, as described in detail elsewhere (C.E. Chamberlain, V. Kraynov, K.M. Hahn, *Methods Enzymology* (In Press 2000)), by taking images of cells containing only GFP-Rac or Alexa-PBD alone, and quantifying the relative intensity of emission in the FRET channel and that in the GFP or Alexa-PBD channel. A broad range of intensities was examined and a line was fit to these for accurate determinations. These corrections had to be applied carefully when studying rapidly moving objects such as ruffles. If the ruffle moved between acquisition of the FRET, GFP, or Alexa images, the subtractive correction process would remove light from the FRET image in the wrong place, generating artefactual FRET localizations. Data from moving features was used only when careful inspection showed the feature to be coincident in the Alexa, GFP, and FRET images, and controls were performed with images taken in different orders. A low pass filter kernel was applied to the corrected FRET image to remove high frequency noise (K. Castleman, *Digital Image Processing* (Prentice Hall, New Jersey, 1996), pp.207-209). Image processing and microscope automation were performed using Invision ISEE software. Images were contrast stretched and formatted for display using Adobe Photoshop software.

We tested Rac and PBD fused to GFP mutants that undergo FRET (ECFP and EYFP) (S. Dharmawardhane, D. Brownson, M. Lennartz, G.M.

Bokoch, *J Leukoc. Biol* 66(3), p. 521- 527 (1999); B. A. Pollok and R. Heim, *Trends Cell. Biol.* 9(2), 57-60 (1999)). Unfortunately, their spectral overlap was more problematic than that of Alexa and GFP, making the corrections described here more difficult. In addition, the GFP mutants showed roughly 25% the FRET of FLAIR. We decided to use FLAIR for these reasons, but the ability to monitor Rac activity simply through protein expression may justify using GFP mutants in some applications.

Simple GFP-Rac fluorescence revealed pools of Rac1 at the nucleus, in the juxtanuclear region, and in small foci throughout the cell prior to stimulation. Confocal and deconvolution imaging showed the nuclear Rac to be concentrated at the nuclear envelope, and expression and immunostaining of HA-tagged Rac1 indicated that these localizations were not an artifact of GFP tagging. Addition of PDGF or serum led to formation of moving ruffles throughout the cell periphery within 2 minutes. These contained GFP-Rac, shown by phalloidin staining of fixed specimens to colocalize with actin (data not shown). The FRET images showed a stark contrast between the level of Rac activation in the ruffles and the nucleus. No FRET was seen at the nucleus despite the high concentration of Rac1 there, while the moving ruffles showed the highest FRET, clearly restricted to the ruffle and discernable from the rest of the cell. The Rac activation remained tightly correlated with the position of the ruffle even as it moved throughout the cell.

As a negative control, FRET was imaged in cells expressing a mutant of GFP-Rho, a close relative of Rac that should not bind PBD. This GFP-Rho Q63L mutant, which generates high levels of GTP-bound protein, produced clear Rho localizations but no corresponding FRET signals (data not shown).

These studies provide the first direct evidence that Rac1 activation is restricted to the site of actin polymerization, independent of the overall distribution of the protein. Rac activation was tightly correlated with moving ruffles, indicating that structures specifically associated with the ruffle were either binding and concentrating activated Rac or that growth-factor induced Rac activation was specifically localized to ruffles. The function of the Rac found at the nuclear envelope remains uncertain. It may be activated for regulation of transcription at times later than those tested here, or may be activated for an unknown role by other stimuli. When activation was concentrated in a small area

such as a ruffle, spatially resolved FRET could detect significant activation changes too small to appreciably alter the overall levels of cellular Rac activity. Our data showed that FRET provided much greater sensitivity and selectivity than following Rac activation simply by imaging Alexa-PBD localization

5 (Figure 9, panels C and D). FRET produces much lower backgrounds and provides complete selectivity even when the biosensor can bind to multiple proteins (i.e. Alexa-PBD also binds to cdc42). Unlike simple localization, FRET can provide quantitative measures of activation levels. It is important to note that the Alexa-PBD could have been sterically hindered from reaching Rac1 in some

10 locations, so that not all activated Rac1 may have been detected. Nonetheless, a FRET signal in a given location does reveal that Rac activation is occurring there.

Rac has been shown to be essential for the directed movement of cells during chemotaxis, and for extension of the front end of cells during motility

15 (C.Y. Chung et al, Proc. Natl. Acad. Sci. U.S.A. 97(10), 5225-5230 (2000). We used FLAIR to ask if Rac activation in polarized, motile cells occurred in particular subcellular localizations to regulate localized actin behaviors. A 'wound' was scraped in a monolayer of confluent Swiss 3T3 fibroblasts, causing cells to become polarized and move into the open space. For wound healing

20 experiments, Swiss 3T3 fibroblasts were induced to undergo polarized movement as previously described (R. DeBiasio, G.R. Bright, L.A. Ernst, A.S. Waggoner, D.L. Taylor, *J. Cell. Biol.* 105, 1613-22 (1987). The cells were cultured in Dulbecco's modified Eagle's medium (GIBCO) supplemented with 10% fetal calf serum at 37°C. Cells were trypsinized and then plated on glass

25 coverslips. They were grown to a confluent monolayer and maintained for an additional 3-4 days. Cells were then wounded by creating a straight laceration with a sterile razor blade. Cells along the edge of the wound were microinjected with 200 micrograms/ml GFP-Rac c-DNA. Six hours after the wound was formed, cells expressing the GFP-Rac were microinjected with 100 micromolar

30 Alexa-PBD and allowed approximately 10 minutes for recovery. Media was then replaced with DPBS containing 10% FCS to reduce background fluorescence. Images were obtained as described above, using exposure times of 1 second for GFP, 1 second for Alexa-PBD, and 5 seconds for FRET. FLAIR revealed highest Rac activation in the juxtannuclear area, and a gradient of Rac

activity highest near the leading edge and tapering off towards the nucleus (Figure 10).

Quantitative analysis clearly indicated that the gradient was correlated with the direction of movement. In individual cells, total Rac activity was measured in two subcellular locations: where activity was highest at the front of the cell, and lowest at the very rear of the cell. These values, measured in squares three microns on a side, were used to calculate the gradient (Percent gradient = $100 \times [\text{front} - \text{back}] / \text{back}$). Of the 16 cells examined, twelve had higher Rac activity at the leading edge, while four showed a slight negative gradient. For cells with high Rac activity at the leading edge, the gradient was $128 \pm 51\%$, while for cells showing the reverse gradient, it was only $9 \pm 4\%$ (mean \pm standard error). The gradient was much broader than the narrow area at the leading edge where actin polymerization occurs (Y.L. Wang et al, *J. Cell Biol* 101, 597-602 (1985; J.A. Theriot and T.J. Mitchison, *Nature* 352, 126-131 (1991). Other activities required for motility, such as depolymerization of fiber networks to recycle monomers and delivery of molecules to the leading edge (O.D. Weiner et al., *Nat. Cell. Biol.* 1, 75-81 (1999) occur throughout the region where Rac is activated. Perhaps Rac is acting over a broader cell area to activate multiple downstream effectors, each producing different effects in more restricted locations. Other studies have shown tight localization of molecules downstream of Rac, at the leading edge or in regions immediately behind it to regulate a variety of functions associated with motility (F. Michiels et al., *Nature* 375, 338-340 (1995).

The prevalence of Rac activation around the nucleus was quantified by scoring 16 cells. All cells showed both juxtanuclear and nuclear GFP fluorescence. Of these, fourteen showed juxtanuclear FRET, and none showed nuclear FRET. It was noteworthy that small areas of the nucleus sometimes showed a FRET signal, but these could be due either to cytoplasmic Rac associated with the nuclear envelope or to juxtanuclear localizations lying over the nucleus. The localization of activation within the juxtanuclear Rac often did not parallel Rac distribution, with "hot spots" of FRET within areas of lower Rac concentration. The meaning of the juxtanuclear Rac localizations is unclear, but their morphology and distribution suggests activation within the ER, golgi. or vesicle populations, perhaps consistent with recent reports suggesting an

important role for Rac in ER to golgi transport (M.P. Quinlan, *Cell Growth Diff.* 10(12), 839-854 (1999), and in pinocytic vesicle cycling (Ridley A.J., Paterson H.F., Johnston C.L., Diekmann D., Hall A., *Cell* 70,401-10 (1992).

In summary, we have described a novel approach to quantify the spatial
5 distribution and rapidly changing levels of Rac signaling in living cells. FLAIR
provided the first direct demonstration that Rac activity is spatially regulated to
generate specific actin behaviors in different subcellular regions. Activated Rac
was tightly coupled to small membrane ruffles even as the ruffles moved
throughout the cytoplasm, yet was broadly distributed as a gradient at the leading
10 edge of motile cells. These very different activation patterns suggest that the cell
will utilize different distributions of activated Rac depending on the number and
localization of downstream targets.

Observations of perinuclear Rac activation and Rac at the nuclear
envelope beg further study. Ultimately, FLAR can reveal how different stimuli
15 interact to affect Rac through the complex circuitry of an intact cell. We have
focused here on spatial control of signaling. The ability of the biosensor to
quantify the level and kinetics of activation should also prove very useful, as
accumulating evidence indicates that Rac and related proteins do not function as
simple binary switches. Different levels and kinetics of activation produce
20 profoundly different results (T. Joneson, *Mol. Cell. Biol.* 19(9) 5892-901(1999).
Using FLAIR together with other biosensors of different wavelengths it should
be possible to examine the balance between Rac, Ras, and other protein
activation levels undergoing rapid changes in real time. With increasing access
to FRET imaging equipment, the technique we describe provides a relatively
25 straightforward way to greatly extend the utility of readily accessible GFP fusion
proteins.

Example 8: Activation biosensors for cdc42.

Like Rac, cdc42 is a member of the Rho family of small GTPases
30 involved in signal transduction in eukaryotic cells. Cdc42 becomes “activated”
by releasing GDP and binding GTP. Such GTPases interact with a host of
downstream effectors, ultimately resulting in one or the other cellular response
via a variety of phosphorylation cascades. In these experiments, a unique
synthetic fluorophore with environment-sensitive fluorescence properties was

linked to the cdc42 binding domain ("CBD") of the Wiscott-Aldrich Syndrome Protein (WASP) to generate a CBD biosensor. Upon binding to the cdc42, this fluorescent CBD biosensor is able to increase its fluorescence intensity by up to 3.5-fold, providing a convenient measure of endogenous cdc42 activation in living cells or for in vitro applications (concept outlined in figure 6b).

Experimental Protocols:

Production of recombinant proteins. DNA encoding the Cdc42-binding fragment of human WASP containing the CRIB motif and surrounding amino acids (WASP amino acids 201 to 321) was amplified by PCR from ATCC clone # 99534. This peptide fragment has the following amino acid sequence (SEQ ID NO:13).

DIQNPDITSSRYRGLPAPGSPADKKRSGKKKISKADIGAPSGFKHVSHV
GWDPQNGFDVNNLDPDLRSLFSRAGISEAQLTDAETSKLIYDFIEDQGGL
EAVRQEMRRQEPLPPPPPS

The full sequence of the WASP protein is as follows (SEQ ID NO:14):

MSGGPMGGRP GGRGAPAVQQ NIPSTLLQDH ENQRLFEMLG RKCLTLATAV
VQLYLALPPG AEHWTKEHCG AVCFVKDNPQ KSYFIRLYGL QAGRLLWEQE
LYSQLVYSTP TPFHTFAGD DCQAGLNFAD EDEAQAFRAL VQEKIQKRNQ
RQSGDRRQLP PPPTPANEER RGGLPPLPLH PGGDQGGPPV GPLSLGLATV
DIQNPDITSS RYRGLPAPGP SPADKKRSGK KKISKADIGA PSGFKHVSHV
GWDPQNGFDV>NNLDPDLRSL FSRAGISEAQ LTDAETSKLI YDFIEDQGGL
EAVRQEMRRQ>EPLPPPPPS RGGNQLPRPP IVGGNKGRSG PLPPVPLGIA
PPPPTPRGPP PPGRGGPPPP PPPATGRSGP LPPPPPGAGG PPMPPPPPPP
PPPPSSGNGP APPPLPPALV PAGGLAPGGG RGALLDQIRQ GIQLNKTPGA
PESSALQPPP QSSEGLVGAL MHVMQKRSRA IHSSDEGEDQ AGDEDEDEW
DD

The DNA fragment encoding SEQ ID NO:13 was subcloned into pET23a (Novagen) as a C-terminal 6His fusion. Site-specific cysteine mutants were constructed by QuikChange (Stratagene) mutagenesis using synthetic oligos and the presence of mutations was confirmed by DNA sequencing. Resultant constructs were transformed into BL21DE3 strain of *E.coli* (Novagen), and the proteins were produced by expression at 30°C for 5 hours in 1 L Leuria-Bertani media (Sigma) in the presence of 100 µg/ml of carbenicilin. Expression was

induced with 0.5 mM IPTG at OD₆₀₀=0.8-1. Cells were collected by centrifugation and stored at -20°C until use.

Cell pellet was resuspended in cold lysis buffer (25 mM Tris-HCl, pH 7.9, 150 mM NaCl, 5 mM MgCl₂, 5% glycerol, 1 mM PMSF, 2 mM β-mercaptoethanol), and briefly sonicated on ice. Lysozyme and DNase were added to the suspension to a final concentration of 0.1 mg/ml and 100 U/ml, respectively, and solution was incubated with occasional stirring at 4°C for 30 min. Lysate was centrifuged (12,000 g, 30 min), and the clarified supernatant was incubated with 1 ml of Talon resin (*Clontech*) at 25°C for 30 min. The resin containing bound CBD was separated from the lysate by brief low-speed centrifugation and washed twice with 15 ml of lysis buffer. Finally, the resin was washed with the lysis buffer, supplemented with 10 mM imidazole, and poured into a column. Elution was performed with 5 ml of the lysis buffer, containing 60 mM imidazole. Fractions containing bulk of CBD (as evidenced by SDS gel) were combined and dialyzed against 1 L of dialysis buffer (25 mM Na₂HPO₄ (pH 7.5), 10 mM NaCl) for 5 hours at 4 C. Solution was then concentrated using Aquacide powder to a final protein concentration of 2 to 10 mg/ml, and dialyzed once again against dialysis buffer. Final preparation was flash-frozen in 100 μL aliquots and stored at -80°C. Generally, 1 to 3 mg of CBD was obtained from 1 L of cell culture. Recombinant 6His-tagged cdc42, RhoA and Rac1 were produced by analogous procedures. The enzymes were determined to be >90% active by the GTP binding assay (Knaus, U.G., Heyworth, P.G., Kinsella, B.T., Curnutte, J.T., and Bokoch, G.M. (1992) *J. Biol. Chem.* 267, 23575-23582).

Conjugation of CBD with fluorescent dyes - Dialyzed CBD sample (100-150 μM) was gently inverted with 6 to 7-fold molar excess of the reactive dye at 25°C for 3 to 4 hours. The reaction was stopped by addition of 10 mM dithiothreitol (DTT), and the mixture was incubated for 15 min. Unreacted dye was separated from the labeled protein using G25-Sepharose (Pharmacia) gel filtration column equilibrated and developed with 25 mM Na₂HPO₄ (pH 7.5). Purity of the eluting fractions was analyzed by running on SDS gel and visualizing fluorescence. Only the fractions containing minimal amounts of free dye were used in the subsequent experiments. Dye-to-protein ratio was determined by measuring CBD concentration ($\epsilon^{280}=8,250 \text{ M}^{-1}$), and A4C

concentration at 617 nm ($\epsilon = 70,000 \text{ M}^{-1}$ in dimethylsulfoxide) or Alexa546 at 554 nm ($\epsilon = 104,000 \text{ M}^{-1}$ in 50 mM potassium phosphate, pH 7.0). Concentration of CBD was independently confirmed by Coomassie Plus assay (Pierce) calibrated with bovine serum albumin as a standard. Dye-to-protein ratios thus obtained varied between 0.8 and 1.2, 1.7 and 2.1 for the single-dye and dual-dye conjugates, respectively. Aliquots of the labeled CBD (15 to 50 μM) were stored at -80°C . No significant loss of binding ability was observed after 6 months of storage. In this example, CBD-conjugates were made with the dye shown in Figure 11, and are referred to as mero-CBD (for merocyanines dye conjugated to CBD).

Analyses of the solvatochromic activation indicator - A solution of mero-CBD (300 nM) in assay buffer was mixed 1:1 (v/v) with solutions of cdc42 (concentration as indicated), pre-equilibrated with 10 mM GDP or GTP γ S as described in Knaus, U.G., Heyworth, P.G., Kinsella, B.T., Curnutte, J.T., and Bokoch, G.M. (1992) *J. Biol. Chem.* 267, 23575-23582. Emission wavelength of 630 nm and excitation wavelength of 600 nm were used to acquire excitation and emission spectra, respectively. For nucleotide dependence, cdc42 (500 nM) was pre-incubated with varying concentrations of GTP γ S (1 to 500 nM). Rac1 and RhoA GTPases were pre-equilibrated with GTP S in the same manner.

Assay of cdc42 activity in stimulated neutrophil cell lysates. - Freshly prepared neutrophils (2.5×10^7 cells/sample) were incubated in a Krebbs-Ringer HEPES buffer supplemented with 5.5 mM glucose (KHRG), in the presence of 1 mM CaCl_2 and 1 μM of fMet-Leu-Phe tripeptide (fMLP) for various periods of time at 37°C (see Benard, V., Bohl, B.P., and Bokoch, G.M. (1999) *J. Biol. Chem.* 274, 13198-13204). Stimulation was stopped by adding equal volume of 2X lysis buffer, supplemented with 2 % Nonidet P-40, 2 $\mu\text{g/ml}$ aprotinin. After a brief vortexing the lysates were clarified by centrifugation and immediately analyzed for cdc42 activity with mero-CBD as described above. In control samples, the lysates were pre-equilibrated with either GDP or GTP γ S at 30°C for 20 to 30 min. This was a convenient method to determine the extent of cdc-42 activation in cell lysates, as shown in Figure 16.

Using the structural information available (Abdul-Manan et al. (1999) *Nature* 399, 379-383; Kim et al. (2000) *Nature* 404, 151-158), we had selected

three positions for placing the fluorescent dye within the CBD peptide (I233, D264 and F271, by WASP numbering), that could experience a considerable change in solvent polarity as a result of binding to cdc42 (see Figure 14). F271 and D264 are located in the loop contacting the effector "switch" domain of the GTPase, whereas I233 appears to interact with the hydrophobic pocket formed by residues at the N-terminus of cdc42 (see Fig. 14). The three CBD residues were mutated to cysteines, easily amenable to covalent modification with the thiol-reactive derivatives of the solvatochromic dyes. Recombinant mutant CBD proteins were overexpressed in bacteria, purified and site-specifically modified with several of solvatochromic dyes. Only the conjugates with the dye shown in Figure 11 are described here.

Among the three mutants tested, the mero-CBD-F271C conjugate exhibited the largest (ca 3.5-fold) fluorescence change in response to binding of activated cdc42 (see Figure 15).

Fluorescence increase as a measure of CBD-cdc42 interaction - The functionality and specificity of mero-CBD was characterized by measuring fluorescence in the presence of saturating amounts of cdc42-GDP or cdc42-GTP γ S. Approximately 3.5-fold increase in both excitation and emission maxima was observed, when the probe was bound to cdc42-GTP γ S, but not cdc42-GDP. Negligible increase (<5%) was also observed in the presence of activated Rac1. No effect on the mero-CBD fluorescence was observed when RhoA-GTP γ S was present. Furthermore, even in the presence of excess of activated Rac1 and RhoA, GDP- and GTP γ S-bound forms of cdc42 were easily distinguished by CBD-A4C fluorescence (data not shown). This result demonstrated the suitability of the CBD indicator for use in live cells where different Rho GTPases may be present at comparable concentrations.

Figure 18 demonstrates use of the biosensor in living cells. To eliminate potential artifacts due to varying cell thickness, uneven illumination, etc., the biosensor was loaded into the cell together with CBD labeled with nonresponsive Alexa546 fluorophore. The ratio of the mero-CBD image to the CBD-Alexa image provided a quantitative measure of the extent of GTPase activation. The warmer colors show areas of higher cdc42 activation.



All publications, patents, and patent documents are incorporated by reference herein, as though individually incorporated by reference. The invention has been described with reference to various specific and preferred embodiments and techniques. However, it should be understood that many variations and modifications can be made while remaining within the spirit and scope of the invention.

09839577-042001